LIST OF ENCLOSURES FOR US EXAMINER

Christian John Cook (assigned to The Horticulture and Food Research Institute of New Zealand Limited)
United States Patent Application No. 09/743,690

	Document	Basis
1.	S R Gallagher (1992) "GUS Protocols: using the GUS gene as reporter of gene expression". Academic Press, San Diego, California, pp 221.	Christeller declaration, page 4.
2.	B W Hicks (2002) "Green fluorescent protein: applications and protocols", Humana Press, Totowa, New Jersey, pp 393.	Christeller declaration, page 4.
3.	D M Mock, P Horowitz (1990), "Fluorometric assay for avidin-biotin interaction". Methods in Enzymology, 184: 234-240. (please note this document replaces that of "M Wilcheck, E A Bayer" referenced in the declaration, which was erroneously identified)	Christeller declaration, page 4.
4.	O Emanuelsson, G von Heijne, (2001), "Prediction of organellar targeting signals" Biochim. Biophys. Acta. Dec 12;1541(1-2):114-9.	Christeller declaration, page 5.
5.	J F Bol, H J M Linthorst, B J C Cornelisson (1990), Annu. Rev. Phytopathol. 28: 113-38.	Christeller declaration, page 22 (Appendix C).
6.	L Frigerio, M de Virgilio, A Prada, F Faoro, A Vitale (1998), "Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. Plant Cell 10: 1031-42.	Christeller declaration, page 22 (Appendix C).
7.	G-Y Jauh, A M Fischer, H D Grimes, C A Ryan, J C Rogers (1998), "d-Tonoplast intrinsic protein defines unique vacuole functions. Proc. Natl. Acad. Sci. 95: 12995-9.	Christeller declaration, page 22 (Appendix C).
8.	N V Raikhel, T A Wilkins (1987), Proc. Natl. Acad. Sci. 84: 6745-9.	Christeller declaration, page 22 (Appendix C).
9.	G Saalbach, R Jung, G Kunze, I Saalbach, K Muntz (1991), "Different legumin domains act as vacuolar targeting signals" Plant Cell 3: 695-708.	Christeller declaration, page 23 (Appendix C).
10.	U Sonnewald, M Brauer, A von Schaewen, M Stitt, L Willmitzer (1991), "Transgenic tobacco plants expressing yeast-derived invertase in either the cytosol, vacuole or apoplast: a powerful tool for studying sucrose metabolism and sink/source interactions. Plant J. 1: 95-106.	Christeller declaration, page 23 (Appendix C).
11.	T A Wilkins, N V Raikhel (1989), "Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos. Plant Cell 1: 541.	Christeller declaration, page 23 (Appendix C).
12.	E A Bayer, T Kulik, R Adar, M Wilchek (1995), "Close similarity among streptavidin-like, biotin-binding proteins from Streptomyces. Biochim Biophysica Acta 1263: 60-66.	Christeller declaration, page 25 (Appendix D).

Document	Basis
13. H Cole, T R Reynolds, J M Lockyer, G A Buck, T Denson, J E Spence, J Hymes, B Wolf (1994), "Human serum biotinidase. cDNA cloning, sequence, and characterization" Biol. Chem. 269: 6566-70.	Christeller declaration, page 25 (Appendix D).
14. K Dakshinamurti, E S Rector (1990), "Monoclonal antibody to biotin". Methods in Enzymology 184: 111-9.	Christeller declaration, page 25 (Appendix D).
15. M Duval, C Job, C Alban, R Douce, D Job (1994), "Developmental patterns of free and protein-bound biotin during maturation and germination of seeds of Pisum sativum: characterization of novel seed-specific biotinylated protein". Biochem. J. 299: 141-50.	Christeller declaration, page 25 (Appendix D).
16. M A Eisenberg, O Prakash, S-C Hsiung (1982), "Purification and properties of the biotin repressor. A bifunctional protein" J. Biol. Chem. 257: 15167-73.	Christeller declaration, page 26 (Appendix D).
17. N M Green (1975), "Avidin", Adv. Protein Chem. 29: 85-133.	Christeller declaration, page 26 (Appendix D).
18. N E Hoffman, E Pichersky, A R Cashmore (1987), "A tomato cDNA encoding a biotin-binding protein" Nucl. Acids res. 15: 3928.	Christeller declaration, page 26 (Appendix D).
19. Y C Hsing, C-H Tsou, T-F Hsu, Z-Y Chen, K-L Hsieh, J-S Hsieh, T-Y Chow (1998), "Tissue- and stage-specific expression of a soybean (Glycine max L.) seed-maturation, biotinylated protein", Plant Biol. 38: 481-90.	Christeller declaration, page 26 (Appendix D).
20. R A Keinanen, M J Wallen, P A Kristo, M O Laukkanen, T A Toimela, M A Helenius, M S Kulomaa (1994) "Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5" Eur. J. Biochem. 220: 615-21.	Christeller declaration, page 26 (Appendix D)
21. J K Korpela, M S Kulomaa, H A Elo, P J Tuohimaa (1981) "Biotin-binding proteins in eggs of oviparous vetebrates. Experentia 37: 1065-6.	Christeller declaration, page 26 (Appendix D).
22. A Krebber, S Bornhauser, J Burmester, A Honegger, J Willuda, H R Bosshard, A Pluckthun (1997) "Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. J. Immunol. Methods, 201: 35-55.	Christeller declaration, pages 26/27 (Appendix D).
23. A T Marttila, K J Airenne, O H Laitinen, T Kulik, E A Bayer, M Wilchek, M S Kulomaa (1998) "Engineering of chicken avidin: a progressive series of reduced charge mutants. FEBS Letters 441: 313-7.	Christeller declaration, page 27 (Appendix D).
24. J Moss, M D Lane (1971), "The biotin-dependent enzymes", Adv Enzymol. Relat.. Areas Mol. Biol. 35: 321-442	Christeller declaration, page 27 (Appendix D).

Document	Basis
25. T Sano, M W Pandori, X Chen, C L Smith, C R Cantor (1995) "Recombinant core streptavidins. A minimum-sized core streptavidin has enhanced structural stability and higher accessability to biotinylated macromolecules" J. Biol. Chem. 270: 28204-9.	Christeller declaration, page 27 (Appendix D).
26. G Tissot, R Douce, C Alban (1997) "Evidence for multiple forms of biotin holocarboxylase synthetase in pea (Pisum sativum) and in Arabidopsis thaliana: subcellular fractionation studies and isolation of cDNA clone" Biochem. J. 323: 179-88.	Christeller declaration, page 28 (Appendix D).
27. H B White, C C Whitehead (1987), "Role of avidin and other biotin binding proteins in the deposition and distribution of biotin in chicken eggs. Discovery of a new biotin-binding protein". Biochem. J. 241: 677-84.	Christeller declaration, page 28 (Appendix D).
28. B Wolf, J Hymes, G S Heard (1990), "Biotinidase". Methods in Enzymology 184: 103-11.	Christeller declaration, page 28 (Appendix D).

GUS Protocols:
Using the GUS Gene
as a Reporter of Gene
Expression

Edited By

Sean R. Gallagher
Hoefer Scientific Instruments
San Francisco, California

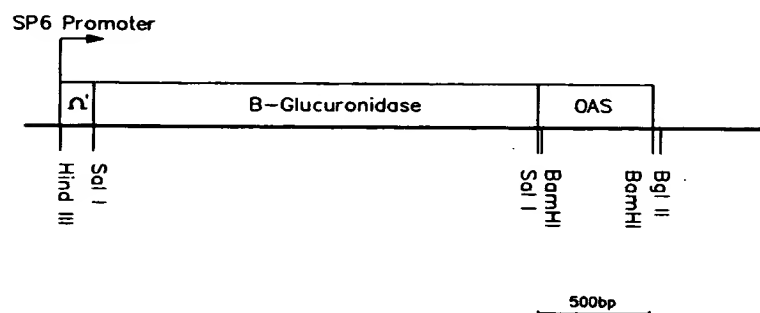**ACADEMIC PRESS, INC.**
Harcourt Brace Jovanovich, Publishers
San Diego   New York   Boston   London   Sydney   Tokyo   Toronto

APPENDIX

B GUS Gene Constructs

Detailed In This Book

SP6 Promoter

Osbourn and Wilson, Chapter 10, Figure 1 Map of SP6 transcription plasmid pJII140 (Sleat *et al.*, 1987). pJII140 was derived from pSP64 (Promega Corp.) via pSP64TMV (Sleat *et al.*, 1986), which contained a *Bam*HI fragment bearing the TMV origin-of-assembly sequence (OAS; genome coordinates 5118-5550). The GUS gene contains no common restriction sites (Jefferson, 1987). *Bgl*II-linearized pJII140 was transcribed as described in the text prior to incubation with TMV coat protein.

210

1kb
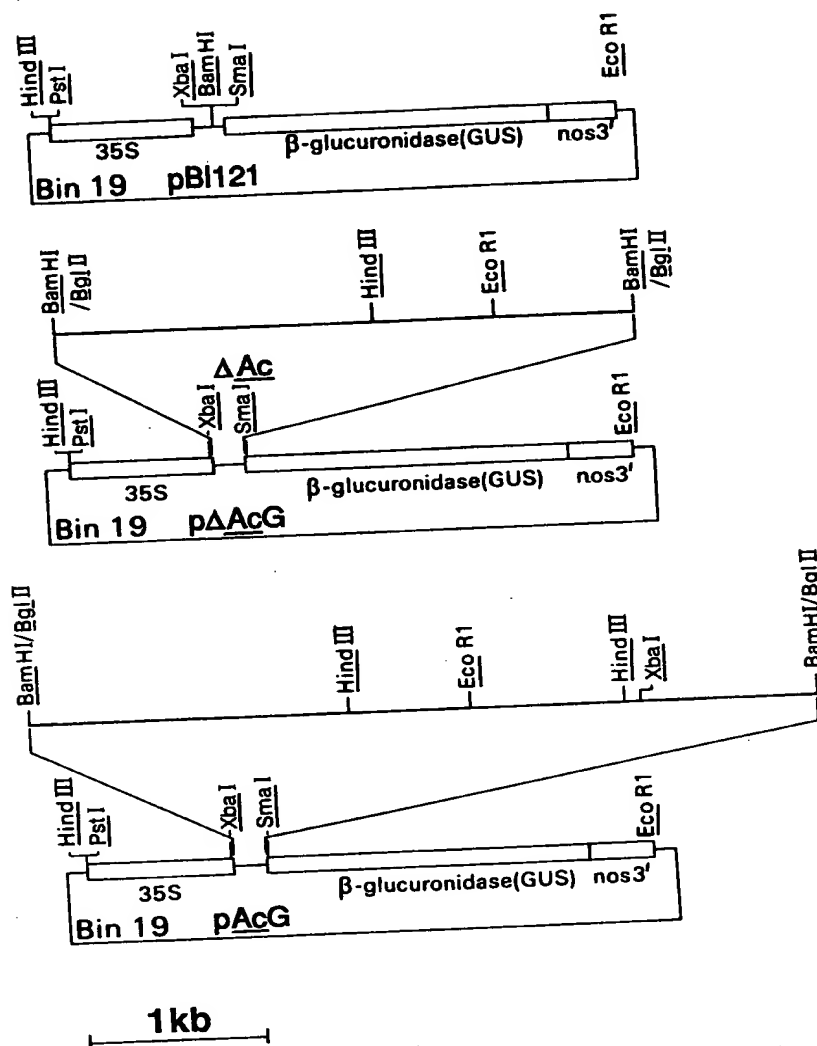
Finnegan, Chapter 11, Figure 2 Restriction maps for the binary vectors pBI121, pΔ*Ac* G, and p*Ac* G. Abbreviations: 35S, cauliflower mosaic virus 35S promoter; GUS, coding region of *E. coli* β-glucuroidase gene; nos3′,3′ termination region from nopaline synthase; Bin 19, binary vector (Bevan, 1984).

		GUS Specific Activity (nmoles/min/mg)	
		CHO	Tobacco
GUS	GpppG $^{17 b}$ $\overset{AUG}{\vert}$ —— GUS —— $\overset{UGA}{\vert}$ $_{67 b}$ GGGUACCGAGCUCGAAUU-3'	1.4	0.01
GUS-(A)$_{25 Dral}$	GpppG $^{17 b}$ $\overset{AUG}{\vert}$ —— GUS —— $\overset{UGA}{\vert}$ $_{67 b}$ AU(A)$_{25}$UUU-3'	20.0	0.34
GUS-(A)$_{50 Dral}$	GpppG $^{17 b}$ $\overset{AUG}{\vert}$ —— GUS —— $\overset{UGA}{\vert}$ $_{67 b}$ AU(A)$_{25}$ GUU(A)$_{25}$ UUU-3'	30.0	0.71

Gallie et al., Chapter 13, Figure 1 GUS constructs used for electroporation of CHO and tobacco cells. Approximately 1 μg of each construct was used for electroporation. See text for details.

Farrell and Beachy, Chapter 9 GUS constructs for protein targeting studies. See pGUSN358 → S in Clontech list below.

Commercially Available

Many of the following plasmids are used in this book, and all are available from Clontech laboratories. Alternatively, plasmids can be obtained by writing to Dr. R. A. Jefferson (see Appendix A) or to the authors of the appropriate work.
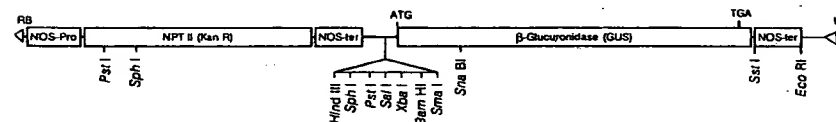

pBI101 (Jefferson et al., 1987) Designed for testing promoter activity, pBI101 confers kanamycin resistance. pBI101 is a derivative of pBIN19 and is unstable unless grown in the presence of kanamycin.

pBI101.2 and pBI101.3 Identical to pBI101, except the reading frames are moved one and two nucleotides, respectively, relative to the polylinker.
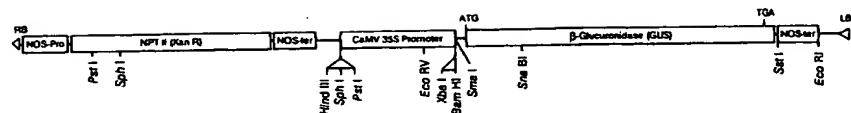
pBI121 (Jefferson et al., 1987)   A derivative of pBI101 containing the 35S promoter of the cauliflower mosaic virus.
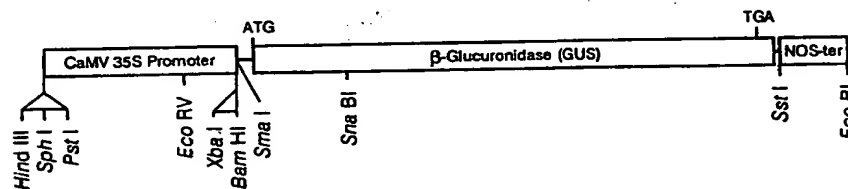
pBI221   The CaMV 35S promoter-GUS-NOS-ter portion of pBI121 was cloned into pUC19 to produce pBI221.

pRAJ255 (Jefferson et al., 1986)   A 1.87 Kb insert containing the GUS gene was cloned into pEMBL9 to produce pRAJ255.

pRAJ260   Similar to pRAJ255 except for a modified Eco R1 site.
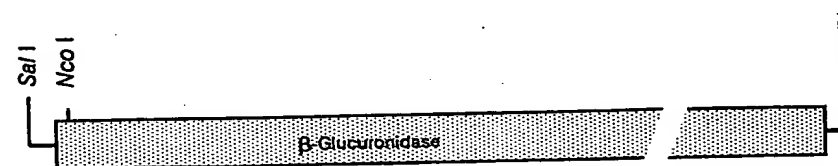
pRAJ275   This derivative of pRAJ255 contains a consensus translational initiator in place of deleted 5' GUS sequences.

pGUSI
GUS g
allows
activity

Refe

Jeffers
E.
8(
Jeffers
β-
pl

Constructs
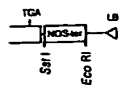
promoter of
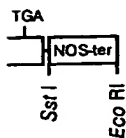
cloned into
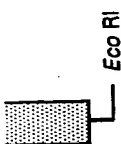
JS gene was
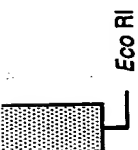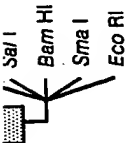
nal initiator in
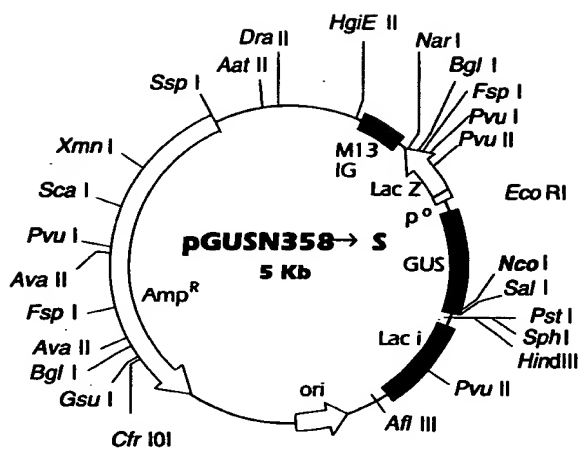
pGUSN358→S (Farrell and Beachy, Chapter 9, this book)   pGUSN358→S contains a GUS gene modified by site directed mutatgenesis to eliminate a glycosylation site. This allows processing by the endoplasmic reticulum without the usual inactivation of GUS activity.

References

Jefferson, R. A., Burgess, S. M., and Hirsh, D. (1986). β-glucuronidase from *Escherichia coli* as a gene fusion marker. *Proc. Natl. Acad. Sci. USA* **86:**8447–8451.

Jefferson, R. A., Kavanagh, T. A., and Bevan, M. W. (1987). GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6:**3901–3907.

Green Fluorescent Protein

Applications and Protocols

Edited by

Barry W. Hicks

Department of Chemistry,
United States Air Force Academy,
Colorado Springs, CO

Humana Press ✳ **Totowa, New Jersey**

19

Green Fluorescent Protein in Transgenic Plants

Brassica Transformation

C. Neal Stewart, Jr., Matthew D. Halfhill, and Reginald J. Millwood

1. Introduction

Until the heterologous expression of *Aequorea victoria* green fluorescent protein (GFP) was demonstrated, scientists working with transgenic organisms had no good alternative to using destructive visible genetic markers. Genes coding luciferase (*1*) and β-glucuronidase (*2*) are the most popular destructive marker genes that have been successfully used in transgenic plants. Although these markers code for sensitive enzymes that have linear dose responses, they require expensive substrates, and are limited to laboratory uses. Most of all, they cannot be used to assay living tissue directly.

GFP offers the possibility to assay vital cellular functions, to determine the transgenic status of plants, and to monitor plant transgene expression in real time, in live cells or intact plants. This chapter focuses on the use of GFP as an enabling biotechnology in the production of transgenic plants, especially *Brassicas*. GFP offers the plant biotechnologist the tool to produce plants in the absence of, or in conjunction with, antibiotic or herbicide markers for selection. It also offers a mechanism to quickly identify transgenic plants in mixed populations. GFP will prove to be an important tool for the making and monitoring of transgenic crops and trees, in the future (*3,4*).

Several GFPs have been shown to be useful in plants. The earliest useful variant was mGFP4, a near-wild-type version that had an altered plant-recognized cryptic intron (*5*). Unfortunately, this GFP was neither bright nor very stable. Improved versions of mGFP4 (mGFP5 and mGFP5-ER) have wild-type chromophores, but have the following mutations: V163A, S175G, and I167T (*5,6*). These mutations confer increased folding at warm temperatures, equal

and dual excitation peaks at 395 and 475 nm, and an emission peak at 509 nm (6). The endoplasmic reticulum version has a signal sequence and HDEL retention signal for targeting GFP to the endoplasmic reticulum. Human codon-optimized S65T mutants have also been useful in plants (7,8). Versions of S65T GFP have a single excitation peak at 489 nm and a red-shifted excitation optimum to (a green) 511 nm (8). Another good choice for plants is the commercially available (Clontech) enhanced GFP, which has the S65T as well as the F64L and Y145F mutations, and is human codon-optimized (9). Other researchers have produced mutants that have been useful in plants (10,11). Recently, GFPs from other organisms have been cloned (12). Plant-optimized GFP, and yellow fluorescent proteins may be expected to be better in plant applications than those currently available. In fact, a priori, *Renilla reniformis* GFP, which has recently been made commercially available by Stratagene, has spectral qualities that should make it brighter in heterologous systems (13). Fluorescent proteins that emit in the yellow and orange spectra have promise in transgenic plant work.

GFP has been used in plant transformation systems as a transformation marker in soybean (14), sugarcane (15), orange (16), tobacco (17), wheat (18), and apple (19), to name a few species. In certain instances, GFP has been used as the sole selectable marker in transgenic plants, demonstrating that a visual marker could be used instead of antibiotic or herbicide selection. Thus far, GFP as the sole selection marker has been proven useful mainly in monocots such as sugarcane (20), barley (21), rice (22), and oats (23). The dicot exception in this case is citrus (16), in which the transformation frequency was compared between GFP-only and GFP plus antibiotic selection. The researchers found that the transformation frequency was the same, but curiously, there were fewer GFP-positive shoots per experiment, using GFP selection (16). One of the benefits of using GFP as the selectable marker is that high-expressing events can be selected very early in the tissue culture and regeneration process.

In this chapter, methods are described that the authors' group has used to transform members of the mustard family (*Brassicaceae*), using GFP-only and GFP in conjunction with antibiotic selection. This lab has produced transgenic *Brassicas* using antibiotic selection (24), and is now using GFP to show proof-of-principle in *Brassica napus*, and also to extend the *Brassica* transformation procedure to a wild relative of the same genus: *Raphanus raphanistrum* (syn. *Brassica kaber*). Various experiments have been performed to demonstrate the efficiencies of GFP-only, or GFP-plus-antibiotic selection. Experiments described here employ a plasmid with GFP and an antibiotic selectable marker, but the goal is to use GFP as the sole selectable marker. Avoiding the use of antibiotic selection could address the criticism of biotechnology opponents who

fear that the horizontal transfer of antibiotic resistance genes could cause medical and ecological emergencies.

2. Materials

1. Surface-sterilized seeds (20% bleach solution for 5 min) from *B. napus* cv Westar.
2. Marashige and Skoog (MS) basal medium (25) for seed (hypocotyl explant source) germination. All plant tissue culture plates are produced using 0.2% Gelrite gellan gum (Sigma, St. Louis, MO) as a gelling agent. All agents are autoclaved, except kanamycin, before media is poured into plates.
3. MS basal medium with 1 mg/L, 2,4-D (MSD1) for 24 h preconditioning hypocotyls, and postco-cultivation.
4. *Agrobacterium tumefaciens* strain GV 3850 containing pBin *mgfp5-er* (35S promoter controlling *mGFPer* gene with linked NOS promoter-controlled *nptII* for kanamycin selection [Fig. 1]).
5. *Agrobacterium* solution (10^8 cells/mL in liquid MS basal medium with acetosyringone 0.05 mM) for co-cultivation with hypocotyls.
6. MSD1 media containing 400 mg/L Timintin to select against *Agrobacterium*, and with or without 20 mg/L kanamycin to select for transformed cells. No kanamycin is used for GFP-only selection.
7. CSRA: MS basal media containing 4 mg/L 6-benzylaminopurine, 2 mg/L zeatin, 5 mg/L silver nitrate, and with or without the above antibiotics to promote organogenesis.
8. CSRB: MS basal media containing 4 mg/L 6-benzylaminopurine, 2 mg/L zeatin, with or without antibiotics.
9. CSE: MS basal medium containing 0.05 mg/L 6-benzylaminopurine plus antibiotics for shoot elongation.
10. MSR: MS basal media containing 0.1% indole butyric acid plus antibiotics to promote root development.
11. 100-nm Petri dishes and GA7 Magenta boxes for tissue culture.
12. Standard dissecting microscope and Spectroline BIB-150 UV lamp.
13. Laminar flow-hood.

3. Methods

3.1. GFP Transformation and Selection in Brassica (24)

1. Seeds are germinated on MS basal media. Zygotic hypocotyls were dissected and chopped into 1-cm-long segments. The hypocotyls segments were placed in a Petri dish containing the *Agrobacterium* inoculum in liquid MS basal medium for 30 min. Periodically shake the segments gently during the 30 min inoculation time Transfer the explants to MSD1 for 1 d, then to MSD1 plus one or no antibiotics (no kanamycin was in the media when using GFP selection only).
2. After 3 d, transfer the tissue to CSRA to initiate shooting. There is a considerable time delay (a few weeks) between shoot initiation and shoot formation using this procedure.
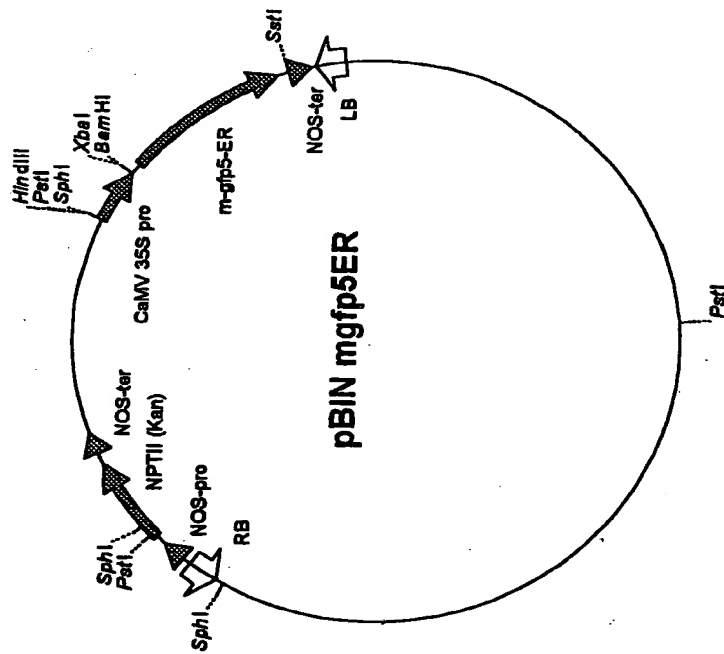
Fig. 1. The binary plasmid, pBin mgfp5ER, which was used for the plant transformation experiments (courtesy of Jim Haseloff). Kanamycin selectable (*nptII*) gene is under the control of the NOS promoter, and the endoplasmic reticulum targeted *GFP* gene is under the control of the 35S promoter from the cauliflower mosaic virus.

3. After another 7 d, (10 d after *Agrobacterium* transformation) transfer the tissue to CSRB. Between 2 and 4 wk GFP fluorescence will appear in calli, then in shoots (*see* **Notes 1–3**).

4. At this point, weekly monitoring with a UV light is required to track transgenic events.

5. When the event callus (fluorescing uniformly green) is approx 0.5 cm in diameter, it is safe to isolate it from the greater tissue and transfer it onto fresh CSRB (*see* **Note 4**). Alternatively, shoots can be transferred to fresh CSRB.

6. Transgenic shoots are transferred to CSE as needed for elongation, then to MSR for rooting.

7. Visually assay for relative transgene expression by comparing GFP emission under UV illumination, thereby allowing selection of the highest-expressing events very early in the transformation process. **Figure 2** shows the product of this method for the transformation of the *Brassica* relative, wild radish (*Raphanus raphanistrum*) on CSRB.

Fig. 2. *Raphanus raphanistrum* hypocotyls segments producing callus stably transformed with mGFP5er under the control of a constitutive promoter. Notice the variation of fluorescence between cut ends. GFP is visualized under UV (365 nm) illumination with no emission filter. (For optimal, color represention please see accompanying CD-ROM.)

4. Notes

1. Much of the success of GFP as an enabling technology in transgenic plants hinges on the success of seeing its production in plants. For lab work, most researchers use epifluorescence microscopes fitted with mercury lamps (~100 W) with blue filters (e.g., 470/40 nm) with 515 nm long-pass emission filters. Of course, without emission filters, one only sees blue reflectance (*see* **refs. 26–28**) for details. In using such arrangements, several researchers have reported background fluorescence that interferes with observing GFP (*14,22,27*). Altering filter choices, such as choosing emission filters of narrower bandwidth, or alternative emission filters should help (*15,21*). Empirical optimization by plant species and tissue types may need to be performed when using blue-light-excited GFPs. The choice of UV-excited GFPs, such as mGFP5, is often ignored as a viable choice by plant scientists. For example, there may be background fluorescence when excited by blue light, but not when excited by UV wavelengths.

2. If one desires to visualize whole plants or organs, then a microscope is not the best tool. For blue-excited GFPs, one can use the photonics of a microscope system, and indeed, Opti-Sciences (Tyngsboro, MA) produces a blue light source with the proper cutoff or bandpass filters for measuring GFP-transgenic plants (GF probe). For UV-excited GFPs, the authors' group and others typically use a portable UV lamp (UVP 100 AP, Upland, CA) with no emission filter, or the lighter Spectroline BIB-150 produced by Spectronics (Westbury, NY). These lamps have a 100 W mercury bulb and a 365-nm filter. The authors group and others have attempted to use less powerful UV lamps with little success. On the other side, we have combined 2–3 of the Spectroline UV lamps, to boost photon excitation irradiation, for more spectacular photographs. To effectively visualize GFP in transgenic plants, the lamp should be very bright and at the proper wavelength. Although the Spectroline or UVP lamps work well for UV excitation of GFP, they would be even more effective if they used a 395 nm filter instead of the 365-nm filter, since the former better matches GFP excitation.

3. UV protective eyewear should be used.

4. There are few tricks to keep in mind when using GFP as a selection for transformation of plants. Tracking transgenic events as early as possible, and keeping the events segregated is desirable. Isolating high-expressing events is important. However, if one excises green fluorescent tissue from the mother explant source, it may die. The authors have been unsuccessful if fluorescent *Brassica* callus is isolated, if the tissue piece is much smaller than 0.5 cm. The UV lamp makes it easy to screen several plates once per week. It also adds the additional benefit of "lighting-up" contaminants that are otherwise hard to see on Petri dishes.

Acknowledgments

References

1. Ow, D. W., Wood, K. V., de Luca, M., deWet, J. R., Helinski, D. R., and Howell, S. H. (1986) Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science* **234**, 856–859.

2. Jefferson, R. A. (1989) The GUS reporter gene system. *Nature* **342**, 837–838.

3. Leffel, S. M., Mabon, S. A., and Stewart, C. N., Jr. (1997) Applications of green fluorescent protein in plants. *BioTechniques* **23**, 912–918.

4. Harper, B. K., Mabon, S. A., Leffel, S. M., Halfhill, M. D., Richards, H. A., Moyer, K. A. et al. (1999) Green fluorescent protein as a marker for expression of a second gene in transgenic plants. *Nat. Biotechnol.* **17**, 1125–1129.

5. Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc. Natl. Acad. Sci. USA* **94**, 2122–2127.

6. Siemering, K. R., Golbik, R., Sever, R., and Haseloff, J. (1996) Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr. Biol.* **6**, 1653–1663.

7. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Improved green fluorescence. *Nature* **373**, 663–664.

8. Chiu, W. L., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H., and Sheen, J. (1996) Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6**, 325–330.

9. Yang, T.-T., Cheng, L., and Kain, S. R. (1996) Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucl. Acid Res.* **24**, 4592–4593.

10. Pang, S.-Z., DeBoer, D. L., Wan, Y., Ye, G., Layton, J. G., Neher, M. K., et al. (1996) An improved green fluorescent protein gene as a vital marker in plants. *Plant Physiol.* **112**, 893–900.

11. Davis, S. J. and Vierstra, R. D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36**, 521–528.

12. Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., et al. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **17**, 969–973.

13. Ward, W. W. (1998) Biochemical and physical properties of green fluorescent protein, in *Green Fluorescent Protein: Properties, Applications, and Protocols* (Chalfie, M. and Kain, S. R., eds.) Wiley and Sons, Chichester, England, pp. 45–75.

14. Ponappa, T., Brzozowski, A. E., and Finer, J. J. (2000) Transient expression and stable transformation of soybean using jellyfish green fluorescent protein (GFP). *Plant Cell Rep.* **19**, 6–12.

15. Elliot, A. R., Campbell, J. A., Dugdale, B., Brettell, R. I. S., and Grof, C. P. L. (1999) Green-fluorescent protein facilitates rapid in vivo detection of genetically transformed plant cells. *Plant Cell Rep.* **18**, 707–714.

16. Ghorbel, R., Juarez, J., Navarro, L., and Pena, L. (1999) Green fluorescent protein as a screenable marker to increase the efficiency of generating transgenic woody fruit plants. *Theor. Appl. Genet.* **99**, 350–358.

17. Molinier, J., Himber, C., and Hahne, G. (2000) . Use of green fluorescent protein for detection of transformed shoots and homozygous offspring. *Plant Cell Rep.* **19**, 219–223.

18. McCormac, A. C., Wu, H., Bao, M., Wang, Y., Xu, R., Elliot, M. C., et al. (1998) The use of visual marker genes as cell-specific reporters of *Agrobacterium*-mediated T-DNA delivery to wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.). *Euphytica* **99**, 17–25.

19. Maximova, S. N., Dandekar, A. M., and Guiltinan, M. J. (1998) Investigation of *Agrobacterium*-mediated transformation of apple using green fluorescent protein: high transient expression and low stable transformation suggest that factors other than T-DNA transfer are rate-limiting. *Plant Mol. Biol.* **37**, 549–559.

20. Elliot, A. R., Campbell, J. A., Brettell, I. S., and Grof, P. L. (1998) *Agrobacterium*-mediated transformation of sugarcane using GFP as a screenable marker. *Aust. J. Plant Physiol.* **25**, 739–743.

21. Ahlandsberg, S., Sathish, P., Sun, C., and Jansson, C. (1999) Green fluorescent protein as a reporter system in the transformation of barley cultivars. *Physiol. Plant.* **107**, 194–200.

22. Vain, P., Worland, B., Kohli, A., Snape, J., and Christou, P. (2000) The green fluorescent protein (GFP) as a vital screenable marker in rice transformation. *Theor. Appl. Genet.* **96**, 164–169.

23. Kaeppler, H. F., Menon, G. K., Skadsen, R. W., Nuutila, A. M., and Carlson, A. R. (2000) Transgenic oat plants via visual selection of cells expressing green fluorescent protein. *Plant Cell Rep.* **19**, 661–666.

24. Stewart, C. N., Jr., Adang, M. J., All, J. N, Raymer, P. L., Ramachsndran, S., and Parrott, W. A. (1996) Insect control and dosage effects in transgenic canola, *Brassica napus* L. (Brassicaceae), containing a synthetic *Bacillus thuringiensis cry IAc* gene. *Plant Physiol.* **112**, 115–120

25. Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plantarum* **15**, 473–497.

26. Ellenberg, J., Lippincott-Schwartz, J., and Presley, J. F. (1998) Two-color green fluorescent protein time-lapse imaging. *BioTechniques* **25**, 838–846.

27. van der Geest, A. H. M., and Petolino, J. F. (1998) . Expression of a modified green fluorescent protein gene in transgenic maize plants and progeny. *Plant Cell Rep.* **17**, 760–764.

28. Rizzuto, R., Carrington, W., and Tuft, R. A. (1998) Digital imaging microscopy of living cells. *Trends Cell Biol.* **8**, 288–292.

V

GREEN FLUORESCENT PROTEIN BIOSENSORS

[25] Fluorometric Assay for Avidin–Biotin Interaction

By DONALD M. MOCK *and* PAUL HOROWITZ

Absolute quantitation of biotin and avidin can be important in many applications of avidin–biotin technology and in studies of biotin nutriture. Many assays for biotin in physiological fluids such as blood and urine are based on the interaction of biotin with avidin and require standardization of the avidin to be used in the assay. Biotin can be quantitated by weight, and then avidin can be quantitated by measuring changes in the optical absorbance spectrum that occur when biotin displaces 4'-hydroxyazobenzene-2-carboxylic acid (HABA) from avidin.[1] An alternate approach is quantitation of avidin by absorbance at 280 nm or by changes in the absorbance spectrum that occur when HABA binds to avidin; biotin can then be quantitated by the optical absorbance changes that occur with displacement of HABA from avidin by biotin.

The interaction of avidin and biotin can also be quantitated by measuring the fluorescence changes that occur when the fluorescent probe 2-anilinonaphthalene-6-sulfonic acid (2,6-ANS) is displaced by biotin from the biotin-binding site on avidin.[2] The probe probably binds to a hydrophobic region at or near the biotin-binding site on avidin. The transition from an aqueous to a hydrophobic environment is associated with a greater than 6-fold increase in the fluorescence intensity, and the wavelength of maximum emission shifts from 463 to 422 nm. When biotin is added, it stoichiometrically displaces the 2,6-ANS (mole ratio 1:1) and reverses the enhancement of fluorescence intensity and the shift in wavelength of maximum emission. This fluorometric method is approximately an order of magnitude more sensitive than the HABA method and is less subject to interference from naturally occurring chromophores; the method is also relatively insensitive to naturally occurring fluorophores. Although this method can be used to standardize either avidin or biotin, for simplicity the standardization of avidin is described in this chapter. The differences when used to standardize biotin are briefly discussed at the end of the chapter.

[1] N. M. Green, *Biochem. J.* **94**, 23c (1965).
[2] D. M. Mock, G. Lankford, D. DuBois, N. Criscimagna, and P. Horowitz, *Anal. Biochem.* **151**, 178 (1985).

Assay Method

Materials. 2,6-ANS is available commercially from Molecular Probes (Junction City, OR). Avidin-D is affinity purified and available commercially from Vector Laboratories (Burlingame, CA). Biotin (*d*-biotin) is from Sigma Chemical Co. (St. Louis, MO). Because the fluorescence of the probe and perhaps even the interaction of the probe and avidin may be sensitive to pH, quantitation should be conducted in an appropriate buffer.[3]

Equipment. We have successfully used two spectrophotofluorometers for this standardization: (1) Farand Mark V spectrophotofluorometers and (2) Shimadzu RF-540 recording spectrophotofluorometers. The excitation wavelength is 328 nm, the emission wavelength 408 nm. A slit width of 5 nm for both excitation and emission beams is suitable.

The wavelength of maximum intensity depends on the solvent system in which the probe is dissolved. Because the equilibrium dissociation constant is about 200 μM, neither the bound nor the free probe concentrations are negligible with respect to the other at the concentrations typically used. Thus, the wavelength of maximum intensity of the probe–avidin complex has been determined by extrapolation to infinite avidin concentration (i.e., all probe bound). The wavelengths are chosen to maximize the difference in fluorescence intensity between bound and free 2,6-ANS based on published maxima for excitation and emission while avoiding scattered light from the excitation beam. In the buffer system used, the true maximum wavelength of emission of free 2,6-ANS is 463 nm and that of bound 2,6-ANS 422 nm. Sensitivity is still satisfactory when monitoring emission at 408 nm; thus, within limits, the choice of excitation and emission wavelengths is not critical.

Titration of 2,6-ANS–Avidin Complex by Biotin

Initially, avidin and 2,6-ANS are added to a quartz cuvette, mixed, and placed in the fluorometer. The maximum fluorescence enhancement is reached in the few seconds required for mixing and insertion of the cuvette; no extra incubation time is required at this point or during the titration. Biotin is then added in increments from a concentrated stock solution, the components are mixed, and the fluorescence intensity is measured after each addition. A typical starting volume is 2.00 ml, and a typical titration would be fifteen 10-μl additions of a 0.12 mM biotin

[3] We chose 0.2 *M N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES Ultrol, Calbiochem, La Jolla, CA), pH 7.0, in 2 *M* NaCl because this buffer gives a satisfactory avidin–biotin interaction in a biotin assay (see D. M. Mock, this volume [24]).

solution in deionized, distilled water. Biotin does not dissolve easily in water at concentrations greater than 2 mg/ml (~8 mM). It may be necessary to make the solution slightly basic with $NaHCO_3$ or NaOH to facilitate dissolving the biotin and then correct back to pH 7.0 with HCl.

The choice of concentration of avidin (or amount of dilution of the unknown avidin source) depends on at least two factors: (1) the sensitivity of the fluorometer and (2) the optical path length of the cuvette. Most fluorescence cuvettes have an optical path length of 1 cm. For experiments requiring high concentrations of 2,6-ANS and thus producing high inner filter effects, we have successfully used a flow cell (Shimadzu RF-540 flow cell, 204038250[4]) with a path length of 2 mm.

When initially setting up the method, it is worthwhile to vary the concentrations of avidin and of 2,6-ANS in the cuvette to achieve the following result when a saturating amount of biotin is added: The decrease in fluorescence intensity should be great enough for accurate quantitation (e.g., ±2% reproducibility) on the available fluorometer. A "saturating" biotin concentration is at least 4-fold greater than the avidin concentration. Typical avidin concentrations are 1–200 μM. Typical 2,6-
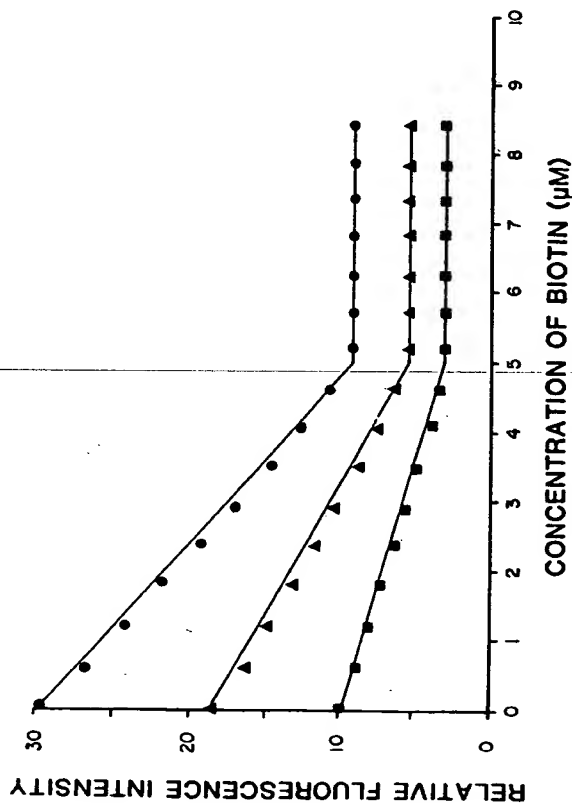
FIG. 1. Titrations of the complex between avidin and 2,6-ANS with biotin at several concentrations of 2,6-ANS. The corrected fluorescence intensity is plotted as a function of the concentration of biotin. The 2,6-ANS concentrations shown are 94.1 μM (top curve), 47.0 μM (middle curve), and 23.5 μM (bottom curve). The avidin concentration is 1.74 μM in all three titrations.

ANS concentrations are roughly equimolar; the point of equivalence (hence, the accuracy of the method) is not dependent on the 2,6-ANS concentration over a fairly wide range (Fig. 1), because the association constant for biotin and avidin is very high ($K_a = 10^{15}$ M^{-1}). For the same reason, the accuracy of this method is not affected by the concentration of avidin in the cuvette (Fig. 2) unless the avidin concentration is too small for accurate quantitation of the fluorescence change or so great that the accompanying high concentrations of 2,6-ANS cause an inner filter effect so large that fluorescence increases cannot be accurately measured and corrected.

The point of equivalence is defined as the point at which all binding sites on avidin are occupied by biotin with no excess free biotin. Because of the high binding affinity, the effects due to equilibrium are negligible at the concentrations used for this method. However, nonlinearity of the
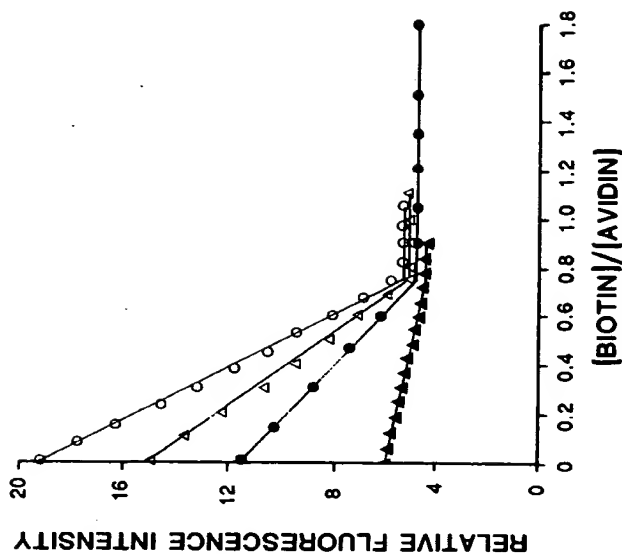
FIG. 2. Titration of the complex between avidin and 2,6-ANS with biotin at several avidin concentrations. The corrected fluorescence intensity is plotted as a function of the ratio of the biotin concentration to the monomeric avidin concentration. The avidin concentrations, determined independently by optical absorbance at 280 nm, are 1, 4, 6, and 8 μM from bottommost curve to topmost. The 2,6-ANS concentration is fixed at 50 μM. The point of equivalence should be 1.0, but this particular batch of avidin consistently demonstrates a value of about 0.75.

titration curve near the equivalence point is sometimes seen. Whether this represents heterogeneity in the avidin preparation or cooperativity in the avidin tetramer (i.e., K_a of one site depends on the number of sites already occupied in that avidin molecule) is not clear.

The point of equivalence can be determined graphically or by linear regression as the point of intersection of the two linear segments of the titration curve; if nonlinearity is seen near the equivalence point, these points should be excluded from the linear regression. The unknown avidin concentration at the point of equivalence $[A]_e$ is related to the biotin concentration at the point of equivalence $[B]_e$ by the following equation:

$$[A]_e = [B]_e/4$$

Keep in mind that the titration may have added significantly to the volume in the cuvette. An alternate equation for the *original* avidin concentration in the cuvette $[A]_0$ (in μmol/liter) is the following:

$$[A]_0 = \text{biotin}/4V_I$$

where biotin is the total micromoles of biotin added at the equivalence point and is equal to the volume of biotin stock added (μl) times the concentration of biotin stock (in μmol/μl), and V_I is the *initial* volume in the cuvette (in liters; e.g., 2.00×10^{-3} liter).

The factor of 4 reflects the assumption that all four sites of tetrameric avidin are available for biotin binding. Using absorbance at 280 nm to standardize avidin independently,[2] we have occasionally found samples of affinity-purified avidin that had three rather than four biotin-binding sites available (Fig. 2). This finding emphasizes the caveat that multimeric binding proteins need only a single functional binding site to be purified by affinity methods. Biotin contamination could result in essentially irreversible occupation of approximately 25% of the sites in a given lot of avidin.

Inner filter effects become important when the number of excitation photons reaching the fluorophores in the region of the cuvette monitored for emission is significantly reduced because of the absorption by fluorophores encountered earlier in the monitored region; at this point, the fluorescence intensity is no longer linear with concentration of fluorophore. Since changes in fluorescence intensity are linear with added biotin in the range of 2,6-ANS concentrations used, the same equivalence point will be determined whether or not the fluorescence intensity is corrected for the inner filter effect. For experiments measuring the binding characteristics of the probe, however, a correction factor according to

the equation of McClure and Edelman[4] is necessary:

$$I_{corr} = I_{obs} \times 2.303 A_{ex}/(1 - 10^{-A_{ex}})$$

where I_{corr} is the corrected fluorescence intensity, I_{obs} the observed fluorescence intensity, and A_{ex} the absorbance of the probe at the excitation wavelength. A_{ex} is calculated as the product of the extinction coefficient at the excitation wavelength times optical path length times concentration of probe.

This fluorescence method will give erroneously low values for avidin if significant amounts of biotin or biotin analogs are present in the avidin sample, provided that the biotin analogs have an intact heterocyclic ring structure and thus bind as tightly to avidin as biotin. Biocytin and biotin methyl ester are examples of such biotin analogs. The effect of biotin analogs with altered ring structures (e.g., dethiobiotin, biotin sulfoxide, diaminobiotin) on this method has not been studied in detail.

Titration of Biotin by Avidin–2,6-ANS Complex

To use the fluorometric method to measure biotin, the source of avidin must first be standardized. With a pure preparation, absorbance at 280 nm can be measured, or the avidin can be quantitated by weight. Alternatively, a primary biotin standard can be used to standardize the avidin, and the avidin, in turn, used to standardize samples of unknown biotin concentration. The titrations to determine the equivalence point are conducted as discussed above if the concentration of biotin in the unknown sample is sufficiently large. If not, a concentrated solution containing avidin and 2,6-ANS can be used to titrate the biotin solution. In this reverse titration, the fluorescence should first increase linearly at a modest slope with sequential additions of avidin–2,6-ANS solution because biotin will immediately displace all the probe into the aqueous phase; at the equivalence point, the slope will increase dramatically because no free biotin remains to displace the probe. Biotin concentration is calculated as follows:

$$[B]_e = [A]_e \times 4$$

This reverse titration (i.e., biotin titrated by avidin) also works if 2,6-ANS is added to the cuvette containing the unknown concentration of biotin rather than along with the avidin. Because the amount of free probe changes little with the avidin additions (except for dilution) until the point

4 W. O. McClure and G. M. Edelman, *Anal. Biochem.* 6, 559 (1967).

of equivalence is reached, the fluorescence intensity increases only after reaching the point of equivalence. Thus, the method of reverse titration may offer a minor advantage in precision compared to titrating with avidin–2,6-ANS solution.

Neither fluorescence enhancement nor reversal is seen with the interaction of 2,6-ANS, streptavidin, and biotin.

Section IV

Applications

Review

Prediction of organellar targeting signals

Olof Emanuelsson, Gunnar von Heijne *

Stockholm Bioinformatics Center, Stockholm University, S-10691 Stockholm, Sweden

Abstract

The subcellular location of a protein is an important characteristic with functional implications, and hence the problem of predicting subcellular localization from the amino acid sequence has received a fair amount of attention from the bioinformatics community. This review attempts to summarize the present state of the art in the field. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Chloroplast; Mitochondria; Presequence; Protein import; Sorting

1. Introduction

The general problem to predict the subcellular location of a protein from its amino acid sequence has long been a central one in bioinformatics. To date, three conceptually different approaches have been proposed: to look for the targeting signals that the cell uses as 'address labels', to base the prediction on the observation that proteins from different cellular compartments tend to differ in subtle ways in their overall amino acid composition, and to use evolutionary relationships (based on the endosymbiotic origin of organelles) to infer the subcellular localization. There are even one or two 'meta-methods' in which outputs from a range of 'primary' prediction/ analysis methods are combined in an optimal way. Each approach has its strengths and weaknesses, and since no across-the-board benchmarking tests have

been performed, it is not yet possible to make a fair comparison between all the different methods proposed by different authors.

In this review, we have chosen to first discuss the most commonly used methods for predicting individual subcellular localizations – the secretory pathway, mitochondria, chloroplasts, and the nucleus – and then describe a couple of attempts to construct integrated predictors that try to 'sort' proteins between multiple compartments. The reader is also referred to a recent (and somewhat more ambitious) review by Nakai [1] for further details.

2. Prediction of signal peptides for secretion

N-Terminal signal peptides target proteins to the secretory pathway in eukaryotic cells, and for translocation across the cytoplasmic membrane in bacteria. It has long been known that they have a tripartite design with a short positively charged amino-terminal segment (n-region), a central hydrophobic segment (c-region), and a more polar C-terminal seg-

* Corresponding author. Fax: +46-8-153679.
E-mail address: gunnar@dbb.su.se (G. von Heijne).

ment that is recognized by the signal peptidase enzyme. The first methods to identify signal peptide sequences were published already in the mid 1980s [2–4], but the currently most widely used method is the neural network-based SignalP predictor [5]. SignalP combines two different neural networks: one that is trained to discriminate between residues that belong and do not belong to a signal peptide (the S-score), and one that is trained only to recognize signal peptidase cleavage sites (the C-score). The cleavage site is predicted by multiplying together the C-score and the negative 'derivative' of the S-score (this serves to focus the prediction on the region where the S-score changes from high to low), while the discrimination between proteins that have and do not have a signal peptide is based on the mean S-score evaluated from the N-terminus to the predicted cleavage site. The current version of SignalP was trained on three different signal peptide data sets – one with eukaryotic signal peptides, one with signal peptides from Gram-negative bacteria, and one from Gram-positive bacteria – and hence is to some extent optimized for different organisms.

SignalP-HMM is a new version of SignalP that is based on a hidden Markov model formalism [6]. This predictor was developed in order to improve the discrimination between signal peptides and N-terminal transmembrane anchor segments, but is in other respects comparable to the original SignalP predictor.

According to a recent benchmarking study [7], SignalP and SignalP-HMM perform equally well when it comes to discriminating between proteins with and without signal peptides, although the neural network version seems to be slightly better in predicting signal peptidase cleavage sites, Table 1. SignalP-HMM is however clearly superior for discriminating between cleavable signal peptides and N-terminal anchors. The two SignalP versions clearly outperformed the other programs tested, and thus seem to be the best signal peptide predictors available at the moment.

It should be mentioned that Chou recently reported a method similar in spirit to a weight-matrix method but including statistics on pairwise correlations between the positions closest to the signal peptidase cleavage site [8]; however, this method was not included in the benchmarking study.

Table 1
Performance of the two versions of SignalP: hidden Markov model version (HMM) and neural network version (NN)

SingalP version	Cleavage site location, % correct			Discrimination, MCC			
				SP/non-SP			SP/SA
	Euk	G−	G+	Euk	G−	G+	Euk
HMM	69.5	81.4	64.5	0.94	0.93	0.96	0.74
NN	72.4	83.4	67.5	0.97	0.89	0.96	0.39

The cleavage site location is measured in the percentage correctly assigned cleavage sites. The discrimination is measured in Mathews' correlation coefficient (MCC) which is one (1) for a perfect prediction and zero (0) for a totally random assignment [33]. The discrimination is given both between signal peptide-containing (SP) and signal peptide-lacking (non-SP) proteins and between secreted proteins (SP) and proteins anchored in the membrane (SA). The table is adapted from [6].

3. Prediction of mitochondrial targeting peptides

Mitochondrial targeting peptides are enriched in positively charged residues (Arg in particular), lack negatively charged residues, and have the ability to form amphiphilic α-helices [9]. The amphiphilic structure is important for binding to receptors in the outer mitochondrial membrane [10,11], and the net positive charge may be needed during the ΔΨ-driven import across the inner mitochondrial membrane [12].

Three popular methods for predicting mitochondrial targeting peptides are TargetP [13], MitoProt [14], and Predotar (see Table 3). Both Predotar and TargetP are neural network predictors and are conceptually similar to SignalP. They are not clear-cut single location predictors since they also deal with other presequences; both investigates chloroplast transit peptide presence and in addition to this TargetP handles signal peptide prediction. Predotar is essentially aimed for plant sequences. The performance of Predotar and TargetP is discussed in Section 6 and summarized in Table 2.

MitoProt predicts localization of a protein by calculating a number of physicochemical parameters from its amino acid sequence, and then computing a linear discriminant function (LDF) which is compared to a cutoff for mitochondrial/non-mitochondrial localization prediction. Both MitoProt and Tar-

Table 2
Comparison of localization prediction for three multi-category predictors

Predictor	Location	Plant set			Non-plant set		
		% Correct	Sensitivity	Specificity	% Correct	Sensitivity	Specificity
TargetP	Chloro.	85.3	0.85	0.69	90.0	–	–
	Mito.		0.82	0.90		0.80	0.67
	Secr.		0.91	0.95		0.96	0.92
PSORT	Chloro.	69.8	0.47	0.69	82.5	–	–
	Mito.		0.66	0.87		0.81	0.60
	Secr.		0.82	0.74		0.64	0.93
Predotar	Chloro.	84.8	0.82	0.77	76.3	–	–
	Mito.		0.86	0.87		0.86	0.50
	Secr.		(0.80)	n/a		(0.65)	n/a

The plant data set contains 940 proteins (from Swiss-Prot release 36) and the non-plant set contains 2738 proteins (from Swiss-Prot release 37) with annotated localization. Note that Predotar is not intended for use on non-plant set, hence its partly poor performance on this set. Sensitivity is the fraction of true positive predictions relative to the set of proteins known to be localized in respective compartment. Specificity is the fraction of true positive predictions relative to the set of proteins predicted to respective compartment. 'Percent correct' refers to the fraction of all proteins in a set for which the correct location is predicted.

getP suggest a potential cleavage site of the predicted mitochondrial targeting peptides.

Predotar, TargetP and MitoProt only predict N-terminal mitochondrial targeting sequences, and no method exists that will identify import signals present elsewhere in the protein, although such signals are known to exist [15,16].

4. Prediction of chloroplast transit peptides

N-Terminal chloroplast transit peptides have highly variable lengths, contain very few negatively charged residues, and are highly enriched for hydroxylated amino acids. Two neural-network based predictors are available: ChloroP [17] and Predotar (see Table 3). ChloroP also includes a separate module (based on a weight matrix) for predicting the

transit peptide cleavage site. A comparison of localization prediction performance between Predotar and TargetP (of which ChloroP is a part) using 940 plant sequences from Swiss-Prot can be found in Table 2.

Many thylakoid proteins have composite targeting signals with a typical transit peptide followed by a thylakoid targeting signal. The latter is usually very similar to the signal peptides found on secretory proteins, and can be identified by SignalP or SignalP-HMM (our unpublished data). A specialized weight matrix for predicting the cleavage site is available for thylakoid signal peptides [18].

5. Prediction of nuclear localization signals

Nuclear localization signals are composed of one (monopartite) or a pair of (bipartite) short positively

Table 3
Web addresses of predictors

Predictor	Web address (URL)
ChloroP	http://www.cbs.dtu.dk/services/ChloroP/
MitoProt	http://www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter
predictNLS	http://maple.bioc.columbia.edu/predictNLS/
Predotar	http://www.inra.fr/Internet/Produits/Predotar/
PSORT	http://psort.nibb.ac.jp/
SignalP	http://www.cbs.dtu.dk/services/SignalP/
TargetP	http://www.cbs.dtu.dk/services/TargetP/
TMHMM	http://www.cbs.dtu.dk/services/TMHMM/

charged stretches in the protein chain. The mono-partite nuclear localization signal has a short consensus sequence, K(K/R)X(K/R), and binds to a pocket on the surface of the importin α receptor [19]. In bipartite nuclear localization signals, the monopartite motif is combined with a second small cluster of basic residues, 10–12 residues N-terminal to the first.

The basic clusters can be found anywhere within the protein chain, and are exposed on the surface of the folded protein. Since the entire chain has to be searched for nuclear localization signals, it is difficult to avoid false positive predictions. The best predictor available at the moment is based on a large collection of mono- and bipartite motifs [20]. It is capable of finding 43% of the known nuclear proteins with no false positive predictions on the set of Swiss-Prot entries (release 38) with unambiguously annotated localization. This is achieved through collection of known NLSs and their homologues, and applying an 'in silico mutagenesis' to extend the motifs as far as possible without matching any non-nuclear proteins.

6. Integrated methods for predicting subcellular localization

In these days of whole-genome sequencing, what is obviously needed are integrated prediction methods that somehow represent the entire protein sorting potential of the cell and assign the most likely subcellular localization to a protein based on its amino acid sequence. This also includes sorting within an organelle or a pathway: between, e.g., the mitochondrial outer membrane, intermembrane space, inner membrane, and matrix, or between the different compartments along the secretory pathway. In eukaryotic cells, the number of distinct compartments is thus very large.

The pioneering work in this area is due to Nakai and Kanehisa [21,22]. His PSORT program now distinguishes between 17 different subcellular localizations (10 for a newer, retrained version called PSORT 2 that uses a slightly different decision algorithm), and integrates a number of pre-existing prediction programs as well as calculated characteristics such as overall amino acid composition within a unified framework [23,24]. Drawid and Gerstein [25]

have recently presented a system that is similar in spirit to PSORT but uses a different formalism (Bayesian statistics) for integrating multiple kinds of information (everything from SignalP predictions to microarray expression profiles). The method was applied to the full *Saccharomyces cerevisiae* proteome, and thus provides estimates of the fraction of all yeast proteins found in different compartments. A predictor based only on overall amino acid composition and pairwise residue correlations has been developed by Chou [26].

The TargetP predictor [13] has a more limited scope than PSORT, and only differentiates between secretory proteins, mitochondrial proteins, chloroplast proteins, and everything else. The method looks for N-terminal sorting signals by feeding the outputs from SignalP, ChloroP, and an analogous mitochondrial predictor (not available as a stand-alone predictor) into a 'decision neural network' that makes the final choice between the different compartments. Although not yet integrated into TargetP, membrane proteins can be predicted with high reliability by programs such as TMHMM [27,28]. TargetP predicts signal peptides with high sensitivity and specificity but performs less well on mitochondrial targeting peptides and chloroplast transit peptides, Table 2. Modules for predicting cleavage sites in the different targeting signals are also included in TargetP; again, performance is much better on the signal peptides than on the other two classes of peptides.

Predotar is primarily aimed at predicting the chloroplast/mitochondrion sorting problem (thus dealing with plant sequences), and can also predict dual localization – both chloroplastic and mitochondrial – which is an existing reality for some proteins [29]. The level of overall prediction accuracy is around 85% on a plant test set, the same as for TargetP, Table 2. The two predictors differ however somewhat in their performance on the subsets and trying both predictors on sequences of interest could prove useful.

Finally, an interesting approach to subcellular localization prediction has been presented by Eisenberg and co-workers [30]. They use a protein's 'phylogenetic profile' (i.e., a list of the presence or absence of orthologs to the query protein in all fully sequenced genomes) to predict its localization,

based on the assumption that the endosymbiont origin of different compartments will be reflected in the phylogenetic profiles of their respective proteomes. Thus, mitochondrial proteins (even the nuclearly encoded ones) will be most highly related to proteins from bacteria such as *Rickettsia prowasekii* [31], whereas chloroplast proteins will be most highly related to those found in photosynthetic bacteria.

Unfortunately, the different methods discussed in this section have not been evaluated together using a common benchmark (since the different methods do not distinguish between the same set of compartments, such an evaluation is not trivial). TargetP has the conceptual advantage that it tries to identify biologically well-characterized sorting signals and hence allows a certain amount of 'critical evaluation by eye' after the prediction has been made. The phylogenetic-profile approach also has a clear biological foundation, and again a human user may critically evaluate the results (i.e., the list of orthologs) against his or her biological knowledge. The purely statistical methods are at a disadvantage in this respect since they are based on sequence characteristics that are not easily evaluated by eye and, insofar as they incorporate amino acid composition measures, only correlate with subcellular localization indirectly (e.g., as a result of surface-exposed residues being adapted to a low-pH environment [32]).

7. Conclusions

The complex compartmentalization of a biological cell cannot yet be accurately captured by bioinformatics. For compartments where the sorting signals can to a good approximation be regarded as short stretches of amino acids with little interaction with the rest of the protein, the sequence analysis tools now available do a decent job. In cases where the sorting signals are presented in the context of a folded protein, however, they are very difficult to identify and one often has to resort to purely statistical approaches (amino acid composition) or methods based on sequence similarity. With improved fold recognition and three-dimensional structure prediction algorithms, it may eventually become possible both to detect these more complex sorting signals

and to predict the location of a protein based on its general surface characteristics. In any event, the prediction of subcellular protein localization will most likely remain an important problem area for bioinformatics for some time to come.

References

[1] K. Nakai, Adv. Protein Chem. 54 (2000) 277–344.
[2] G. von Heijne, Eur. J. Biochem. 133 (1983) 17–21.
[3] G. von Heijne, Nucleic Acids Res. 14 (1986) 4683–4690.
[4] D.J. McGeoch, Virus Res. 3 (1985) 271–286.
[5] H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Protein Eng. 10 (1997) 1–6.
[6] H. Nielsen, A. Krogh, Intell. Syst. Mol. Biol. 6 (1998) 122–130.
[7] K.M.L. Menne, H. Hermjakob, R. Apweiler, Bioinformatics 16 (2000) 741–742.
[8] K.-C. Chou, Protein Eng. 14 (2001) 75–79.
[9] G. von Heijne, EMBO J. 5 (1986) 1335–1342.
[10] Y. Abe, T. Shodai, T. Muto, K. Mihara, H. Torii, S. Nishikawa, T. Endo, D. Kohda, Cell 100 (2000) 551–560.
[11] T. Muto, T. Obita, Y. Abe, T. Shodai, T. Endo, D. Kohda, J. Mol. Biol. 306 (2001) 137–143.
[12] W. Voos, H. Martin, T. Krimmer, N. Pfanner, Biochim. Biophys. Acta 1422 (1999) 235–254.
[13] O. Emanuelsson, H. Nielsen, S. Brunak, G. von Heijne, J. Mol. Biol. 300 (2000) 1005–1016.
[14] M.G. Claros, P. Vincens, Eur. J. Biochem. 241 (1996) 779–786.
[15] N. Wiedemann, N. Pfanner, M.T. Ryan, EMBO J. 20 (2001) 951–960.
[16] C.M. Lee, J. Sedman, W. Neupert, R.A. Stuart, J. Biol. Chem. 274 (1999) 20937–20942.
[17] O. Emanuelsson, H. Nielsen, G. von Heijne, Protein Sci. 8 (1999) 978–984.
[18] C.J. Howe, T.P. Wallace, Nucleic Acids Res. 18 (1990) 3417.
[19] M. Hodel, A. Corbett, A. Hodel, J. Biol. Chem. 276 (2001) 1317–1325.
[20] M. Cokol, R. Nair, B. Rost, EMBO Rep. 1 (2000) 411–415.
[21] K. Nakai, M. Kanehisa, Proteins Struct. Funct. Genet. 11 (1991) 95–110.
[22] K. Nakai, M. Kanehisa, Genomics 14 (1992) 897–911.
[23] P. Horton, K. Nakai, in: Fifth International Conference on Intelligent Systems for Molecular Biology (ISMB-97) Halkidiki, Greece,, 1997, pp. 147–152.
[24] K. Nakai, P. Horton, Trends Biochem. Sci. 24 (1999) 34–35.
[25] A. Drawid, M. Gerstein, J. Mol. Biol. 301 (2000) 1059–1075.
[26] K.-C. Chou, Proteins Struct. Funct. Genet. 43 (2001) 246–255.
[27] A. Krogh, B. Larsson, G. von Heijne, E. Sonnhammer, J. Mol. Biol. 305 (2001) 567–580.
[28] S. Möller, M. Croning, R. Apweiler, Bioinformatics 17 (2001) 646–653.

[29] I. Small, H. Wintz, K. Akashi, H. Mireau, Plant Mol. Biol. 38 (1998) 265–277.

[30] E. Marcotte, I. Xenarios, A. van Der Bliek, D. Eisenberg, Proc. Natl. Acad. Sci. USA 97 (2000) 12115–12120.

[31] S.G.E. Andersson, A. Zomorodipour, J.O. Andersson, T. SicheritzPonten, U.C.M. Alsmark, R.M. Podowski, A.K. Näslund, A.S. Eriksson, H.H. Winkler, C.G. Kurland, Nature 396 (1998) 133–140.

[32] M. Andrade, S. O'Donoghue, B. Rost, J. Mol. Biol. 276 (1998) 517–525.

[33] B. Matthews, Biochim. Biophys. Acta 405 (1975) 442–451.

PLANT PATHOGENESIS-RELATED PROTEINS INDUCED BY VIRUS INFECTION

J. F. Bol and H. J. M. Linthorst

Department of Biochemistry, Gorlaeus Laboratories, Leiden University, Einsteinweg 5, 2333 CC Leiden, The Netherlands

B. J. C. Cornelissen

MOGEN International, Einsteinweg 97, 2333 CB Leiden, The Netherlands

INTRODUCTION

The hypersensitive response triggered by an incompatible plant-pathogen interaction is accompanied by numerous metabolic changes (for reviews see 5, 8, 20, 37, 119). One general effect is the induction of enzymes that are involved in the synthesis of ethylene, a plant hormone able to induce many stress responses. The induction of enzymes from the phenylpropanoid pathway, oxidative enzymes, and hydroxyproline-rich glycoproteins (HRGP, extensin) is partially related to cell wall modifications that lead to the formation of a physical barrier to further spread of the pathogen from the site of infection. Peroxidases are involved in the polymerization of alcohol derivatives of aromatic compounds such as coumaric, ferulic, and sinapic acid into lignin and suberin and the cross-linking of these polyphenols to extensin molecules from the cell wall matrix. In addition to inducing a defense aimed at localizing the pathogen, various antimicrobial compounds are synthesized. These compounds include phytoalexins synthesized along branches of the phenylpropanoid pathway leading to furano-coumarins or isoflavonoids,

113

Table 1 PR proteins induced in Samsun tobacco (NN genotype) by TMV infection

Group	Acidic PR proteins Name[a]	Mol wt (kd)	Basic PR proteins Name	Mol wt (kd)	Function
1	1a	15.8	16 kd	16.0	Unknown
	1b	15.5			
	1c	15.6			
2a	2	39.7	Gluc.b	33.0	β-1, 3-Glucanase
	N	40.0			
	O	40.6			
	Q'	36.0			
2b	O'	25.0			β-1, 3-Glucanase
3	P	27.5	Ch.32	32.0	Chitinase
	Q	28.5	Ch.34	34.0	
4	s1	14.5			Unknown
	r1	14.5			
	s2	13.0			
	r2	13.0			
5a	R	24.0	Osmotin	24.0	Unknown (TL-proteins[b])
	S	24.0			
5b			45 kd	45.0	Unknown

[a] Nomenclature is according to Fritig et al (38)
[b] Thaumatin-like proteins

hydrolytic enzymes that may attack polysaccharide components from cell walls of the pathogen, and proteinase inhibitors that specifically block the digestive enzymes of herbivores (see Ryan, this volume). Collectively, these induced reactions are thought to contribute to an active defense mechanism of the plant known as "induced resistance."

Similar metabolic changes are induced by the incompatible interaction of a plant with either a fungus, a bacterium, or a virus (20, 37, 119). In this chapter we focus on the "pathogenesis-related" (PR) proteins that are induced in plants after infection with necrotizing viruses. In the early 1970s, Van Loon & Van Kammen (121) and Gianinazzi et al (42) independently reported the de novo synthesis of several proteins in tobacco plants reacting hypersensitively to infection with tobacco mosaic virus (TMV). These proteins were characterized by their acidic nature (43, 115), their resistance to proteases (117), and their extracellular location (88). More recently, basic homologues to a number of acidic PR proteins have been identified. In tobacco, the basic PR proteins are not secreted into the intercellular space of the leaf but accumulate in the vacuoles. Table 1 lists over 20 acidic and basic PR proteins purified to homogeneity from TMV-infected Samsun NN tobacco by Fritig et al (38, 53, 55). A number of acidic PR proteins have also been purified by Antoniw et al (2), Pierpoint (94) and Van Loon et al (120). Different nomenclatures are used by various research groups (6); here, we adhere to that used by Fritig et al (38).

The tobacco PR proteins listed in Table 1 have been classified into five groups (120). Group 1 consists of the PR-1 proteins of unknown function. A serological cross-reactivity between the acidic and basic PR-1 proteins has been reported (55). Group 2 contains PR proteins with β-1,3-glucanase activity, which was detected first for PRs 2, N and O (56) and later for Q' and O' (38). The TMV-induced basic tobacco β-1,3-glucanase (Gluc. b) is probably identical to a similar enzyme produced in tobacco tissue cultured on auxin-containing medium (33, 104). PRs 2, N, O, Q' and Gluc. b are serologically related but antisera to these proteins do not cross-react with PR-O' (38). PR-O', which occurs as a dimer, is less strongly induced after TMV infection than the other PR proteins. In an in vitro assay, the enzymatic activities of PR-O and Gluc. b have are similar whereas that of the other β-1,3-glucanases are 20- to 100-fold lower (38, 56).

TMV-infection was found to induce four tobacco chitinases, namely the acidic PR proteins P and Q and the basic proteins Ch. 32 and Ch. 34 (66). The basic enzymes had been identified earlier in cultured tobacco tissue (103). All four proteins are serologically related and constitute group 3 (Table 1). Although the basic chitinases are six times more active than their acidic isozymes in an in vitro assay, the more acidic forms (PRs P and Q) were estimated to account for one third of the TMV-induced chitinase activity in

tobacco (66). Chitinases and β-1,3-glucanases are thought to play a role in a plant's defense mechanism against fungal infection (102).

The low molecular weight proteins classified as group 4 have been characterized in less detail. Two TMV-induced proteins with molecular weights of approximately 13 kd and 15 kd were recognized by Pierpoint (94) and Van Loon et al (120). Fritig et al (38) and Kauffmann (55) purified two 13.0-kd proteins (s2 and r2) and two 14.5-kd proteins (s1 and r1); r1 and r2 comigrate, as do s1 and s2, in alkaline nondenaturing polyacrylamide gels. The serological relationship between proteins of groups 2a and 4 reported by Van Loon et al (120) was not found by Kauffmann (55).

Two acidic 24-kd proteins are induced after TMV infection of *Nicotiana tabacum* L., cv Samsun NN (55) or *Nicotiana tabacum* L. cv Xanthi nc (95) tobacco; these are named PR-R and PR-S. In addition, a TMV-induced basic 24-kd protein has been identified that is serologically related to PRs R and S

Figure 1 Induction of PR-mRNAs in Samsun tobacco (NN genotype) by treatment with salicylate or infection with TMV. Seven Northern blots were loaded with RNA from healthy plants (lanes H), plants sprayed with 5 mM salicylate (lanes S) and plants infected with TMV (lanes T). The blots were hybridized to ^{32}P-labeled cDNA clones corresponding to the acidic (upper panels) and basic (lower panels) isoforms of PR-1 proteins, β-1,3-glucanases and chitinases, and the acidic isoform of PR-R/S. The estimated size of the mRNAs (kb) is indicated in the right margin of each blot. Reprinted, with permission of J. F. Bol et al (7).

(55). This latter protein is probably identical to osmotin, a protein induced by salt stress of tobacco (105). All three proteins, placed in group 5a (Table 1), show a 65% amino acid sequence homology to the sweet-tasting protein thaumatin (23, 95, 105). A similar degree of homology was found with a maize protein that is a bifunctional inhibitor of α-amylase and proteases of insects (97), suggesting that the tobacco proteins play a role in plant defense against insects. However, purified PR-R and PR-S did not inhibit α-amylase or various proteases in an in vitro assay (B. Fritig, personal communication).

The purification of a TMV-induced 45-kd tobacco protein has recently been reported (55) and, because of its serological cross-reactivity to PR-R and PR-S, has tentatively been placed in group 5b. Furthermore, a serological relationship between PR-R and PR-S and PR proteins of group 1 has been reported (120), but this relationship is not paralleled by a detectable amino acid sequence similarity.

Proteins related by serology or amino acid sequence similarity to the TMV-induced tobacco proteins from groups 1, 2a, 3, and 5a occur in over 20 plant species, including monocots and dicots (5, 6, 118). This indicates that these proteins, which may constitute up to 10% of the total soluble protein under various stress conditions, are highly conserved in the plant kingdom. Below, the current state of knowledge about these proteins and their corresponding genes is reviewed.

INDUCTION OF PR PROTEINS

Similar patterns of proteins are induced after the incompatible interaction of plants with different types of pathogens and each interaction, in turn, induces a broad range of defense reactions. For example, after infection with TMV, tobacco plants become resistant to infection from unrelated viruses, fungi, and bacteria (40). Moreover, induction of PR proteins by abiotic elicitors generally results in acquired resistance to pathogens, with very few exceptions (36). Some well-known chemical inducers of PR proteins include polyacrylic acid, ethephon (which metabolizes into ethylene), aromatic compounds such as benzoic acid, salicylic acid and acetyl salicylic acid; amino acid derivatives; the antiviral agents 2-thiouracil and dioxohexahydrotriazine; the herbicide phosphinotricin; and barium and manganese salts (5, 118; J. F. Bol et al, unpublished results). It has been claimed that PR proteins are also induced by elicitor preparations extracted from fungal or bacterial cultures (65, 75). Injection of tobacco with a xylanase solution resulted in an induction of PR proteins, suggesting that released oligosaccharide fragments may act as signal molecules (70). On the other hand, spraying of tobacco plants with a glucan preparation from the mycelial walls of *Phytophthora megasperma* did not induce PR proteins or enzymes from the phenylpropanoid pathway (38).
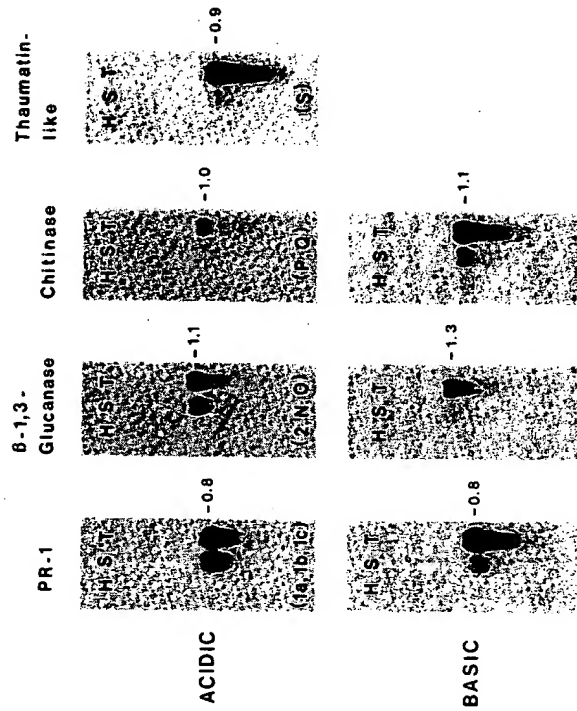
Similarly, floatation of tobacco leaf discs on such an elicitor solution did not induce PR proteins (L. C. van Loon, personal communication).

Figure 1 shows seven Northern blots that reveal the accumulation of mRNAs encoding acidic and basic PR proteins in the leaves of healthy tobacco plants (lanes H), plants sprayed with a solution of 5 mM salicylate (lanes S), and plants inoculated with TMV (lane T). All mRNAs are present at a relatively low concentration in the leaves of healthy plants and are strongly induced after TMV infection. In addition, a number of mRNAs are induced by salicylate, notably those encoding acidic and basic PR-1 proteins, acidic glucanases, and basic chitinases. An analysis of a cDNA library made to poly(A)-RNA from TMV-infected tobacco yielded clones corresponding to several induced mRNAs that do not encode known PR proteins (51). Figure 2 is a Northern blot showing the induction by salicylate and TMV-infection of mRNA, called "A", encoding a 40-kd protein of unknown function, and of mRNA encoding a glycine-rich protein (GRP) (113).

For six of the mRNAs shown in Figures 1 and 2, accumulation reaches a maximum concentration about four days after inoculation with TMV in
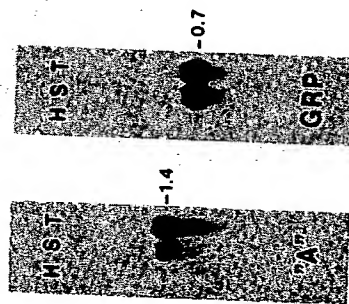
Table 2 Accumulation of TMV-inducible mRNAs in tobacco Samsun (NN genotype).

mRNA	Virus infected leaves	Ethephon sprayed leaves	Healthy plants				
			leaves	roots	stems	flowers	fruits
Acidic PR PR-1	+ᵃ	-	-	-	-	-	-
Gluc.	+	-	-	-	-	-	-
Chit.	+	-	-	-	-	-	-
PR-S	+	-	-	-	-	-	-
Basic PR PR-1	+	+	-	+	+/-	+/-	-
Gluc.	+	+	-	+	+/-	+/-	+/-
Chit.	+	+	-	+	+/-	+/-	-
Non PR A	+	+	-	-	-	-	-
GRP	+	+	-	-	-	-	+

ᵃUnder conditions indicated by a plus sign the accumulation of mRNA is at least 1000-fold higher than under conditions indicated by a minus sign. For experimental details see Memelink et al (80).

and differ in their tissue-specific expression, indicates that the promoters of these genes are controlled by complex regulatory elements. The recent cloning of a number of PR genes is a first step toward characterizing these elements.

PR-1 PROTEINS

cDNAs to tobacco mRNAs encoding acidic PR-1 proteins have been cloned by several groups (24, 27, 76, 92). Three types of mRNAs were identified that could be correlated, respectively, with PR-1a, -1b, and 1c by determining the partial amino acid sequences of these proteins (73, 91). The amino acid sequence indentity between the three acidic proteins is over 90%. The 138-amino-acid-long mature proteins are derived from primary translation products by removal of an N-terminal signal peptide of 30 amino acids. The acidic PR-1 proteins from *Nicotiana tabacum* cvs. Samsun NN, Wisconsin-38, and Xanthi nc show minor sequence variations. Analysis of the cDNA clones revealed that alternate polyadenylation signals are used for each PR-1 mRNA.

Screening of a cDNA library made to poly(A)-RNA from TMV-infected tobacco yielded one clone encoding a basic isoform of the acidic PR-1 proteins (25). The amino acid sequence similarity between the acidic and basic PR-1 proteins is about 67%. The 19-kd basic protein is made with a C-terminal extension of 36 amino acids not present in the acid isoforms. A genomic clone has been identified encoding another basic PR-1 protein that contains a C-terminal extension of 18 amino acids (89).

Figure 2 Induction of "A" and GRP mRNAs in Samsun tobacco (NN genotype) by salicylate treatment or TMV infection. Northern blots were loaded with RNA from healthy plants (lanes H), plants sprayed with 5 mM salicylate (lanes S) and plants infected with TMV (lanes T). The blots were hybridized to ^{32}P-labeled "A" and GRP cDNA clones.

primary infected leaves. After eight days the mRNAs are detectable in virus-free, systemically induced leaves (51).

The expression of a number of PR genes is developmentally regulated. Chitinase, β-1,3-glucanase, and the corresponding mRNAs are present at high concentrations in roots of healthy plants (34, 103). Memelink et al (80) showed by Northern blot analysis that genes encoding basic PR-1 proteins—basic β-1,3-glucanases, and basic chitinases—are expressed at different levels in roots, stems, and flowers of healthy plants (Table 2). It has been reported that pricking leaves with needles moistened with ethephon, a treatment that causes necrotic lesion-like spots on the leaves, results in induction of genes encoding acidic PR proteins (116). Spraying the leaves with an ethephon solution did not induce the acidic PR proteins but efficiently "turned on" the genes encoding basic PRs and the non-PR proteins A and GRP (Table 2). Similarly, the genes encoding basic β-1,3-glucanases and basic chitinases are coordinately induced by exposure of plants to ethylene (11, 127).

The plus and minus signs in Table 2 represent crude estimates of the various mRNA levels. A more careful serological analysis revealed an organ-specific expression of several acidic PR proteins in different parts of the flowers of healthy tobacco (71). Acidic PR-1 proteins accumulate in sepals, whereas a glycosylated vacuolar form of acidic β-1,3-glucanase is found in pistils. Interestingly, within the class of acidic chitinases, the genes encoding PR proteins P and Q are differentially expressed in sepals (P and Q), pedicels (P only), anthers (low levels of P and Q), and ovaries (low levels of P only). Stress-induced expression of tobacco GRP genes is leaf-specific and occurs in a light-dependent manner (80). The induction of acidic PR proteins by various chemicals in tobacco leaf material is light-dependent (3). The observation that PR genes differ in their response to salicylic acid or ethylene

Infection of tomato plants with either viroids, viruses, or *Cladosporium fulvum* results in the accumulation of a basic 14-kd protein (isoelectric point of 10:7), referred to as p14 or P1. Determination of the primary structure of p14 revealed a 60% amino acid sequence similarity with acid and basic PR-1 proteins of tobacco (73). Antisera to the tobacco and tomato proteins cross-react. Moreover, these antisera permitted the detection of serologically related inducible proteins in cowpea, potato, *Solanum demissum*, *Gomphrena globosa*, *Chenopodium amaranticolor*, maize, and barley (85, 130).

Because of their relative resistance to proteases, PR proteins are well adapted to survive in the intracellular space of the leaf or in vacuoles. Their long half-lives, estimated to be 40–70 hr, permit the accumulation of high levels of PR proteins. In viroid-infected tomato plants p14 was specifically degraded by a constitutive extracellular endoproteinase with a molecular weight of 37 kd (100); possibly, this enzyme plays a role in the turnover of PR proteins.

GLUCANASES AND CHITINASES

General Occurrence

The acidic or basic β-1,3-glucanases or chitinases that are coordinately induced in plants by ethylene, fungal elicitor, fungal or viral infection are all endo-hydrolases. Chitinases catalyze the hydrolosis of β-1,4N-acetylglucosamine linkages of chitin polymers. Most of the purified plant chitinases tested have lysozyme activity, and some enzymes under optimal conditions hydrolyze bacterial peptidoglycan faster than chitin (12). Chitin is not present in healthy plants but may constitute a significant fraction of fungal cell walls, ranging from 0.3% in *Phytophthora* spp. to about 50% in *Fusarium* spp. The other major component in the cell wall of these microorganisms in β-1,3-glucan. Combinations of purified pea chitinase and β-1,3-glucanase inhibited growth of 15 out of 18 fungi tested, whereas the enzymes alone inhibited only one of these fungi (77). Possibly, these hydrolases play a role in the TMV-induced resistance of tobacco to subsequent infection by fungi and bacteria (40).

The combination of enzymes with β-1,3-glucanase and chitinase activity has been found so far in tobacco, potato, tomato, bean, soybean, pea, barley, and maize (12, 84). In addition, chitinases have been identified in ten other plant species including cucumber, melon, and carrot. Below we discuss some of the recently characterized systems.

Tobacco

When tobacco leaf tissue was subcultured on an auxin-containing medium, up to 10% of the soluble protein fraction was found to consist of a 33-kd protein that was identified as a basic β-1,3-glucanase (33). cDNA cloning indicated the induction under these conditions of at least three different mRNAs that encode highly homologous basic β-1,3-glucanases (104). A similar cDNA clone was used in Figure 1 to reveal the induction of these mRNAs by TMV-infection. The basic β-1,3-glucanase is produced as a 359-residue preproenzyme. In a first maturation step the N-terminal signal peptide of 21 residues is removed and the C-terminal sequence becomes glycosylated. Subsequently, a C-terminal sequence of 22 residues with the oligosaccharide side chain is lost (104). Most acidic tobacco β-1,3-glucanases are produced without this C-terminal extension. From an analysis of cDNA clones (H. J. M. Linthorst & J. F. Bol, unpublished observations) and partial protein sequence data, it can be concluded that at least five acidic β-1,3-glucanases with highly similar amino acid sequences are induced in tobacco leaves by TMV-infection. The amino acid sequence identity between basic and acidic β-1,3-glucanases is about 65%.

Analysis of cDNA and genomic clones revealed the sequence of a hormonally regulated β-glucanase from *Nicotiana plumbaginifolia* (28). This enzyme contains a C-terminal extension similar to that of the basic tobacco β-1,3-glucanase. The amino acid sequence identity between the two enzymes is 73%.

Sequencing of cDNA clones corresponding to basic chitinases (103) and acidic chitinases (50; H. J. M. Linthorst & J. F. Bol, unpublished observations) revealed that these PR proteins also are synthesized as precursors with a N-terminal signal peptide. The amino acid sequence identity between these two isoforms is about 65%. The higher molecular weight of the basic chitinases compared to the acidic chitinases (Table 1) is due to a N-terminal extension of 46 amino acids and a C-terminal extension of 6 amino acids (H. J. M. Linthorst & J. F. Bol, unpublished observations). Interestingly, on 40 residues the N-terminal of the basic chitinases show significant sequence similarities to the N-terminal domains of other chitin-binding proteins such as hevein, wheat-germ agglutinin, and rice and nettle lectins (12, 74, 109). The absence of this domain in acidic tobacco chitinases demonstrates that it is not involved in catalytic activity.

Potato

After inoculation of leaves with zoospores of *Phytophthora infestans* or treatment of the leaves with an elicitor derived from *P. infestans* culture filtrate and administered through the petioles, potato plants start to accumulate massive amounts of at least nine major proteins having molecular weights in the range of 10–40 kd (60). Two proteins with molecular weights of 36 and 36.2 kd and isoelectric (pI) points of 9.6 and 9.8, respectively, were shown to be basic β-1,3-glucanases. These proteins were serologically related to β-1,3-

glucanases from bean (see below). Six other proteins with molecular weights between 32.6 and 38.7 kd proved to be basic chitinases with pI values above 7. All six chitinases showed strong cross-reactivity with a bean anti-chitinase antiserum.

cDNA was produced from a potato mRNA that encodes a basic chitinase with a predicted pI of 10.55. After removal of the signal peptide the molecular weight of the mature form is 32.4 kd (39). The enzyme shows considerable sequence similarity to the tobacco chitinases.

Bean

Exposure of bean seedlings to ethylene results in the rapid and coordinate induction of a 32.5-kd chitinase and a 36-kd β-1,3-glucanase with pI values of 9.0 and 9.7, respectively (127). The use of cDNA clones to the mRNAs of chitinase (17) and β-1,3-glucanase (127) as probes, showed that induction of these PR proteins is regulated at the level of transcription. A similar rapid induction of chitinase mRNA was observed in bean hypocotyls in response to wounding or infection with *Colletotrichum lindemuthianum* and after treatment of bean protoplast cultures with fungal cell-wall elicitors (45). The cDNA sequence revealed that the mature chitinase of 301 residues is derived from a precursor with a 27-residue amino-terminal signal peptide. The amino acid sequence identity between the bean chitinase and the basic tobacco chitinase is 73% (103). The polypeptide translated in vitro from bean β-1,3-glucanase mRNA was 4 kd larger than the mature protein, which indicates that this enzyme also is derived from a precursor.

Pea

Two basic chitinases and two basic β-1,3-glucanases have been identified in pea. The accumulation of these enzymes is differentially regulated. One chitinase (Ch2, mol wt 36.2 kd) and β-1,3-glucanase (G1, mol wt 33.5 kd) are produced in pea pods during the course of maturation. Another chitinase (Ch1, mol wt 33.1 kd) and β-1,3-glucanase (G2, mol wt 34.3 kd) are strongly induced in immature pea pods in response to various stress conditions such as inoculation with strains of *Fusarium solani* pathogenic or nonpathogenic to pea, wounding, or treatment with chitosan or ethylene (77). The developmentally regulated and stress-induced isozymes differ in a number of enzymatic properties, which indicates that they fulfill different functions.

Cucumber

Infection of cucumber with tobacco necrosis virus or treatment with salicylate results in the induction of an acidic chitinase with molecular weight of 28 kd. The sequence derived from a cDNA clone shows that the protein of 267 residues matures from a precursor with a N-terminal signal peptide of 25 amino acids (81). The amino acid sequences of the known chitinases from tobacco, potato, bean, and barley are closely related to each other, but no detectable similarity is observable between these enzymes and the cucumber chitinase. However, this latter enzyme shows striking homology to the partial amino acid sequence deduced from a lysozyme/chitinase of *Parthenociccus quinquifolia*.

Barley

Two classes of basic β-glucan endohydrolases in barley have been characterized at the molecular level, each represented by two isozymes. The first class contains β-1,3-1,4-glucanases (isozymes EI and EII) that are primarily responsible for the degradation of cell walls of the nonliving starchy endosperm during germination of barley grains. The second class consists of β-1,3-glucanases (isozymes GI and GII) that also accumulate to relatively high concentrations in barley grains although their substrate in germinating barley is very limited. The observation that after infection with *Erysiphe graminis* β-1,3-glucanase mRNA accumulates to high concentrations in resistant plants but not in susceptible plants, indicates that GI and GII are involved in defense reactions (48). cDNA to EII (35) and GII (48) has been cloned and sequenced. Both proteins are basic polypeptides of 306 amino acids, derived from their respective precursors by removal of the N-terminal signal peptide. The amino acid sequence identity between EII and GII is about 50% and a similar homology is observed with the basic tobacco β-1,3-glucanase.

Another chitinase, a 36-kd endochitinase, is secreted from cells in the aleurone layer during seed germination. cDNA cloning of the mRNA of this chitinase revealed an extensive homology to the tobacco and bean chitinase and also to a 28-kd endochitinase present in barley flour (110).

Maize

Eight proteins, all having molecular weights in the range of 14.2–34.5 kd, were found to accumulate in maize leaves after treatment with mercuric chloride or infection with brome mosaic virus. These were named PRm 1 to 8 (84), and could be divided into at least three classes. PRm 2 (16.5 kd) was serologically related to the tobacco PR-1b protein. Four of these proteins (PRm 3, 4, 5, and 7) were endochitinases. PRm 3 and 4 both have a molecular weight of 25 kd and are serologically related. PRm 5 and 7 have molecular weights of 29 kd and 34.5 kd, respectively; they are serologically related and cross-reacted with an antiserum to the tobacco chitinase PR-P, but are not serologically related to PRm 3 and 4. Finally, PRm 6a (32 kd) and PRm 6b (30.5 kd) appeared to be serologically related β-1,3-glucanases.

large quantities in the intercellular fluid. By using immunogold-labeling techniques, p14 was detectable in the intercellular spaces as well as in association with inclusion bodies within the vacuoles of viroid-infected tomato plants, which suggests a dual localization (124). The basic thaumatin-like NP24 protein of tomato was found to accumulate intracellularly (59).

In potato all basic chitinases and β-1,3-glucanases were predominantly found in the intercellular fluid, which suggests that they are secreted after fungal infection or elicitor treatment (60). Similarly, the basic chitinases and β-1,3-glucanases of barley are secreted by cells in the barley aleurone layer (48, 110). In contrast, cell fractionation and immunocytochemical studies on ethylene-treated bean leaves revealed that all of the basic bean chitinase and most of the basic bean β-1,3-glucanase accumulated in vacuoles (14, 78, 127). Only a small amount of the bean β-1,3-glucanase, but no chitinase, is secreted into the apoplast. Like the acidic tobacco chitinase, the acidic chitinase of cucumber is secreted almost exclusively into the intercellular space of infected leaves (13).

All known PR proteins are synthesized as precursors with an N-terminal signal peptide. Litte homology is observed between these signal sequences, except for their hydrophobic nature and the presence of charged residues at the preferred positions (129). Possibly, they function only in the translocation of the PR polypeptide chains across the membrane of the endoplasmatic reticulum, and signals in the mature proteins determine whether they are transported from the Golgi complex to the vacuoles or are secreted through the plasma membrane. The current view is that a positive signal near the N-terminus of a mature protein specifies its routing to the vacuoles, whereas the absence of such a signal results in a default pathway leading to secretion (for references see 125, 126). This would mean that the C-terminal extension present in the vacuolar basic PR proteins of tobacco, but absent in the extracellular acidic isoforms, has no role in the targeting of these proteins. Similarly, the presence of an N-terminal sequence of 46 amino acids in the basic tobacco chitinase, which is absent in the acidic tobacco chitinase, is not responsible per se for the vacuolar localization of the basic isoform because a similar N-terminal sequence is present in the extracellular basic potato chitinase. Possibly, a modulation of a putative signal near the N-terminus of mature PR proteins determines whether the protein is transported to the vacuoles, is secreted to the apoplast, or whether both pathways are followed. The identification of signals involved in targeting of PR proteins requires further experimentation.

Only two examples of glycosylation of PR proteins are known. The C-terminal extension of the basic tobacco β-1,3-glucanase is glycosylated but the oligosaccharide side chain is not present in the mature protein (104). Moreover, a glycosylated vacuolar form of acidic β-1,3-glucanase has been

THAUMATIN-LIKE PROTEINS

Pierpoint et al (95) found two closely related acidic 24-kd proteins, named R-minor and R-major, that occur in a 2:3 ratio in TMV-infected *Nicotiana tabacum* L. cv Xanthi nc tobacco. These proteins probably correspond to PR-R and PR-S, respectively, in Fritig's (38) nomenclature. cDNA has been cloned, respectively, to the mRNAs of PR-R (23) and PR-S (90). Salt stress of tobacco cells induces two related basic 24-kd proteins, named osmotin I and II, that occur in an approximate ratio of 2:3 (105). Determination of the N-terminal amino acid sequence revealed a 62% homology with PR-R/S. cDNA cloning of a salt-induced basic 24-kd tomato protein (NP24) revealed that this protein is over 90% homologous to the tobacco osmotin (59). The acidic and basic tobacco proteins and the tomato NP24 all show approximately 60% amino acid sequence homology to the sweet-tasting protein thaumatin, which accumulates in the fruits of *Thaumatococcus daniellii*, and to a bifunctional amylase/protease-inhibitor from maize (23, 97). The acidic and basic tobacco proteins do not taste sweet. The basic thaumatin is synthesized as a precursor with an N-terminal signal peptide and a C-terminal extension of six amino acids that are removed during maturation. The tobacco and tomato thaumatin-like proteins also have signal peptides. The basic tomato protein (and possibly also the basic tobacco protein) has a C-terminal extension of 19 residues that is absent in the acidic tobacco protein.

TARGETING OF PR PROTEINS

In tobacco, acidic and basic PR proteins appear to be strictly compartmentalized. The former accumulate predominantly in the extracellular space of the leaf, i.e. the apoplast (88), which has been analyzed in most detail by cell fractionation techniques for acidic PR-1 proteins (86) and β-1,3-glucanases (112). The extracellular location of acidic PR-1 was verified by immunofluorescence microscopy and immunogold-labeling techniques (19, 30, 52). On the other hand, according to available evidence, all basic tobacco PR proteins are sequestered in vacuoles. Immunocytochemical detection revealed that osmotin is concentrated in dense inclusion bodies within the vacuole. Using cell fractionation techniques, the basic tobacco β-1,3-glucanase (112) and chitinase (unpublished observation quoted in 55) were also shown to accumulate in vacuoles. In tobacco, the extracellular acidic PR proteins may serve as a first line of defense against invading pathogens, with the vacuolar basic proteins stored in the vacuole as a second line of defense (9).

Although the basic PR-1 protein in tobacco is not secreted in detectable amounts, its equivalent basic p14 protein in tomato is detectable in relatively

function in suberization and lignification, respectively. After TMV infection, two of the moderate cationic peroxidases are formed in inoculated and systemically induced tobacco leaves (62, 117). The molecular cloning of an anionic lignin-forming peroxidase from tobacco (61) and suberization-associated anionic peroxidases from potato (99) and tomato (98) has been reported. The enzymes show a 40–50% sequence homology to horseradish and turnip peroxidase.

Hydroxyproline-Rich and Glycine-Rich Proteins

Hydroxyproline-rich glycoproteins (HRGP, extensin) make up 1–10% of the wall matrix. Intermolecular cross-linking of extensin monomers by peroxidases results in a network that serves as a matrix for the linkage of polysaccharides (pectin, cellulose) or polyphenols (lignin, suberin) (22). Synthesis of HRGP mRNA is induced by ethylene, fungal infection, elicitors (31, 64), and by infection of tobacco with TMV (79).

Certain plants contain little if any HRGP; instead their cell walls are rich in glycine. A wound-inducible glycine-rich protein (GRP) from petunia was suggested to be a cell wall component (21). Association of a wound-inducible GRP with the vascular system of bean has been confirmed by immunocytochemical techniques (57). Like the TMV-inducible tobacco GRP shown in Figure 2, the petunia and bean GRP are synthesized with N-terminal signal peptides. Interestingly, the sequences of the signal peptides of the tobacco GRP and a light-induced GRP from *Chenopodium rubrum* are highly similar (54, 113). Induction of GRPs by water stress or abscisic acid has been observed in maize and rice, but these are probably cytosolic proteins (44, 83).

Active Defense Proteins

Several other inducible plant proteins besides the group of PR proteins are involved in defense against pathogenic attack. Low molecular weight proteinase inhibitors may constitute up to 5% of the protein in seeds and tubers of Gramineae, Leguminosae and Solanaceae. Their systemic induction in tomato and potato leaves by wounding is thought to be mediated by oligosaccharides released from plant-cell walls by an endopolygalacturonase (101; Ryan, this volume). These proteins inhibit serine endopeptidases that are found in both animals and microorganisms but only rarely in plants. Because of this specificity they are thought to be involved in defending the plant against herbivorous insects.

Pathogens and various stress conditions trigger the synthesis of thionins in the leaves of barley, maize, wheat, oats, and several dicotyledonous plants. These proteins are incorporated in the cell wall and display significant antifungal activity in in vitro assays (4).

TMV-infected protoplasts from hypersensitive tobacco cultivars have been

identified in flower parts of healthy tobacco (71). Available evidence indicates that glycosylation contributes to the stability of a protein but has no role in its targeting (126). Nevertheless, it is interesting to note the similarity in processing of basic tobacco β-1,3-glucanase and the vacuolar wheat germ agglutinin (WGA). Also, WGA is glycosylated at a C-terminal extension of 15 amino acid concomitantly with the removal of a signal peptide and this extension is cleaved off during transport to the vacuole (96).

OTHER VIRUS-INDUCED PROTEINS

Biosynthesis of Aromatic Compounds

The hypersensitive response of plants to pathogens is accompanied by increased activity of enzymes involved in the biosynthesis of aromatic compounds based on the phenylpropane skeleton. These compounds include the cell wall structural polymer lignin, flavonoid pigments and UV protectants, furanocoumarin and isoflavonoid phytoalexins, and wound protectant hydroxycinnamic acid esters. The general pathway in phenylpropanoid metabolism involves the conversion of phenylalanine to 4-coumaroyl-CoA in three steps catalyzed by phenylalanine-amonia-lyase (PAL), cinnamate-4-hydroxylase (C4H), and 4-coumaroyl-CoA ligase (4CL). From this route different branches lead to the synthesis of furanocoumarins, (iso)flavonoids, stilbenoids, suberin, lignin, and other phenolics. Chalcone synthase (CHS) and chalcone flavanone isomerase (CHI) are the key enzymes in the branch toward (iso)flavonoid synthesis. Cinnamoyl-CoA-reductase and cinnamoyl-alcohol-dehydrogenase (CAD) generate the lignin precursor alcohols 4-coumaryl, coniferyl, and sinapyl, which are polymerized by peroxidases to give the three-dimensional lignin network. Recently, the following genes involved in the biosynthesis of these aromatic compounds have been cloned and characterized: PAL from bean and parsley (26, 69), 4CL from parsley (72), CHS from parsley, soybean, *Antirrhinum majus*, *zea mays*, arabidopsis, petunia and bean (32, 46, 131), CHI from petunia (122), and CAD from bean (128). Most of these genes correspond to multigene families that are differentially regulated. The induction of PAL, 4CL, CHS, CHI, and CAD genes by infection, elicitor, and ethylene has been demonstrated (20, 31, 64, 117). Initial experiments on the characterization of regulatory elements in the promoter sequences of these genes have been performed (29, 63).

Peroxidases

At least 12 peroxidase isozymes divisable into three subgroups have been identified in tobacco (61). The function of the vacuolar cationic peroxidases (pI, 8.1–11) is as yet unclear. The moderate anionic (pI, 4.5–6.5) and the anionic (pI, 3.5–4.0) peroxidases are cell-wall associated and are thought to

reported to secrete an inhibitor of virus replication (IVR) into the medium. Virus replication was inhibited in protoplasts and leaf disks and also in intact leaves when IVR was applied by cut stems or by spray. IVR was identified as a 23-kd protein that is not only released by infected protoplasts but also accumulates in the intercellular space of TMV-infected tobacco (108). IVR is not serologically related to AVF, a putative antiviral factor consisting of a 22-kd phosphoglycoprotein (82). A possible serological cross-reaction of IVR with well-characterized tobacco PR proteins has not been analyzed.

Other Inducible Proteins

Differential screening of a cDNA library from cultured parsley cells yielded 18 independent cDNA families of elicitor-induced mRNAs (106). The structure and sequence of a gene corresponding to one of these classes of parsley mRNAs, named PR1, has been deduced (107). This parsley PR1 gene does not show sequence similarity to other known PR genes. Moreover, the 16.5-kd parsley PR1 protein (isoelectric point of 4.7) is synthesized without a signal peptide, suggesting that it accumulates in the cytosol. The PR1 mRNA is induced by fungal infection of parsley.

In addition to p14, a 69-kd PR protein, referred to as P-69, is induced in tomato by citrus exocortis viroid infection. Recently, it was shown that P-69 is an alkaline endoproteinase (pI, 8.5 to 9.0) with optimal activity at pH 8.5 (123). It has been suggested that P-69 is involved in degenerative processes occurring in the chloroplasts of infected plants.

Treatment of Nicotiana plumbaginifolia with ethylene or salicylic acid, or infection of this plant with Pseudomonas syringae results in the induction of the enzyme manganese superoxide dismutase. cDNA cloning revealed that the enzyme is synthesized with an N-terminal leader sequence resembling a transit peptide for mitochondrial targeting. The function of the enzyme could be the dismutation of superoxide radicals generated by the enhanced respiratory oxidation of sugars when stress is applied to the plant (15).

REGULATION OF PR GENE EXPRESSION

PR-1 and GRP Genes

Among the salicylate-inducible tobacco genes shown in Figures 1 and 2, the PR-1 and GRP genes have been characterized in most detail. The gene families encoding acidic PR-1, basic PR-1, and GRP from Samsun NN tobacco each consist of about eight members. After TMV infection, three genes encoding acidic PR-1 proteins (PR-1a, -1b and -1c) are expressed and at least two genes encoding basic PR-1 proteins and four genes encoding GRP (25, 89, 92; H. J. M. Linthorst & J. F. Bol, unpublished observation). The PR-1a and -1c genes are derived from Nicotiana Sylvestris whereas the PR-1b

gene is derived from the other parent of Samsun NN tobacco, N. tomentosiformis (41). The sequence of the PR-1a gene and several pseudo genes from the PR-1 family has been reported (25, 87, 91, 93).

One gene encoding a basic PR-1 protein has been cloned (89). There is limited homology with the PR-1a gene throughout the 5'-flanking sequences with no obvious conserved region. The cloned gene encoding basic PR-1 is not coordinately expressed with genes encoding acidic PR-1, in contrast to the gene corresponding to the basic PR-1 cDNA clone used to probe the blot of Figure 1. This suggests that genes encoding basic PR-1 proteins are regulated differently.

Two genomic GRP clones have been sequenced (113). One gene, named GRP-8, contains an intron of 555 bp and is expressed after TMV infection, whereas the other gene (GRP-4) is not expressed after TMV infection and contains an intron of 1954 bp. A defective splice donor site was thought to be responsible for the "silent" nature of the GRP-4 gene.

The upstream sequences of the PR-1a and GRP genes contain a number of direct and inverted repeats that could have regulatory functions. Most prominent is a 64-bp inverted repeat in the GRP upstream sequence that is almost identical to a 64-bp inverted repeat found upstream of the rubisco small subunit gene of tobacco (113). Upstream sequences of the PR-1a and GRP-8 genes have been fused to reporter genes, and these constructs were used to transform tobacco. In the transgenic plants, 689 bp of the PR-1a upstream sequence and 645 bp of the GRP upstream sequence were sufficient to induce expression of the reporter gene when the plant was sprayed with 5 mM salicylate or infected with TMV (M. D. Van de Rhee, J. A. L. van Kan, M. T. González-Jaén & J. F. Bol, unpublished observations). Deletion studies revealed that sequences located between positions -689 and -643 in the PR-1a promoter and positions -645 and -400 in the GRP-8 promoter were required for induction of reporter-gene expression by the two treatments. The only detectable similarity between the two promoter regions was a small sequence resembling the SV40 core enhancer sequence. Circumstantial evidence suggested that such an enhancer sequence could work in concert with salicylate-responsive elements located near the TATA-box (M. D. Van de Rhee, J. A. L. van Kan, M. T. González-Jaén & J. F. Bol, unpublished observation). Similar fusions of the promoter region of the gene for the bean cell-wall glycine-rich protein with a reporter gene revealed that a 494-bp upstream sequence was sufficient for correct tissue-specific expression and wound-inducibility of the gene (58).

Glucanase and Chitinase Genes

The genome of Samsun NN tobacco contains about eight genes encoding acidic β-1,3-glucanases and a similar number encoding basic β-1,3-

glucanases. Most genes for acidic isozymes are expressed after TMV infection (H. J. M. Linthorst & J. F. Bol, unpublished observation) while at least three of the genes for basic isozymes are active (104). A gene encoding the β-glucanase of Nicotiana plumbaginifolia has been cloned (28); it contains a 669-bp intron near the 5'-end of the open reading frame in the region encoding the N-terminal signal peptide.

The acidic and basic chitinases of tobacco are each encoded by two to four genes, most of which are expressed after TMV infection (50). The structure of these genes has not been reported yet. At least two of the three chitinase genes known to be present in the bean genome are subject to regulation by ethylene. Comparison of the upstream sequences of chitinase genes CH5B and CH5A revealed two conserved regions (16). Region I is located between -927 and -1470 in the CH5B gene and shows 96% homology to the sequence between -533 and -1200 of the CH5A gene. Region II consists of two stretches of 13 bp and 43 bp, respectively, that are perfectly conserved between both chitinase genes. In CH5B these segments occur at -181 to -223 and -252 to -264, whereas in CH5A they reside further upstream. The significance of these homologies has been studied in transgenic plants transformed with CH5B-promoter/GUS fusions (16). Deletion of the upstream sequence to -575 resulted in no significant loss in the induction by ethylene of the reporter gene, indicating that region I is not necessary for this regulation. Deletion to -422 resulted in a 20-fold reduction in ethylene-inducibility but a significant level of ethylene-inducibility remained until the upstream sequences were deleted to -195. These results point to the location of an activator between -575 and -422 and the location of ethylene-responsive elements between -422 and -195 in the CH5B gene. This latter sequence overlaps with the conserved region II. It will be interesting to learn whether similar sequences are found upstream of glucanase genes that are coordinately induced with chitinase genes by ethylene. Small homologies have been observed in the promoter regions of the ethylene-inducible genes encoding PAL, 4 CL and CHS, but their significance has not yet been determined (29, 63).

PR-R/S Genes

Screening of a tobacco genomic library (114) yielded two classes of clones containing reading frames that were colinear with cDNA made to mRNAs (genes) encoding PR-R (23) and PR-S (90). The upstream sequences of the two genes show a large degree of homology over a length of at least 1 kb. A comparison of the upstream sequences of the tobacco PR-1a, GRP, and PR-S genes revealed limited homologies only in the first 100 bp upstream of the transcription initiation site. The location of TMV-responsive elements in the PR-R/S promoters is currently being analyzed.

CONSTITUTIVE EXPRESSION OF PR GENES

Constitutive expression of PR proteins in healthy plants could shed light on the putative role of these proteins in defense mechanisms. The observation that a hybrid of Nicotiana glutinosa x Nicotiana debneyi showed both a constitutive expression of a PR-1 protein and a high level of resistance to TMV is evidence for a role of PR-1 in an antiviral defense (1). Low levels of PR-1 gene expression have been observed in transgenic tobacco plants of the NN genotype and transformed with the coat protein gene of TMV but not in transgenic of the nn genotype tobacco plants. As the transgenic NN and nn plants are both resistant to TMV infection, PR-1 proteins probably are not involved in coat protein–mediated protection (18). Transformation of tobacco with the VI gene of cauliflower mosaic virus (CaMV) causes growth suppression and the development of necrotic spots on stems and leaves of the plants. This phenotype was accompanied by the expression of high levels of PR-1 proteins and an eightfold increase in β-1,3-glucanase activity (111).

Treatment of tobacco with salicylate results in more than 90% inhibition of the replication of both necrotic and systemic viruses. Studies with protoplasts showed that salicylate treatment blocked viral RNA replication without having a detectable effect on host metabolism (49). To test the possibility that salicylate-induced host proteins are involved in this inhibition, tobacco has been transformed with PR-1a and GRP-coding sequences fused to the CaMV 35S promoter, and the PR-S gene has been expressed constitutively in transgenic tobacco (68). Selective expression of PR-1a, GRP, and PR-S was achieved to concentrations comparable to those in TMV-infected plants without concomitant induction of other PR-genes. These transgenic plants were just as susceptible as nontransformed control plants to either localized infection with TMV or systemic infection with alfalfa mosaic virus (68). Moreover, there was no difference between the transgenic plants and the controls in their susceptibility to attack by the beet armyworm or tobacco budworm (10, 67). This indicates that PR-1a, GRP, and PR-S alone are not effective in protecting plants from virus infection or insect attack. Probably, the sequence homology between PR-S and a maize α-amylase/protease inhibitor does not reflect a functional equivalence. Engineering of a cowpea trypsin inhibitor into tobacco results in a resistance to the tobacco budworm (47).

Recently it was shown that expression of a PR-S gene in the minus-strand orientation in transgenic tobacco plants resulted in an 80–90% reduction of expression of the endogenous PR-R/S genes after TMV infection (10). If the same approach could effectively block the induction of hydrolytic enzymes around TMV-lesions in tobacco leaves, one could answer the question whether putative oligosaccharides released by these enzymes from damaged

cell walls constitute the signal(s) that trigger(s) the induction of defense genes at a distance from the site of infection (see Ryan, this volume).

CONCLUDING REMARKS

The host proteins induced by the hypersensitive response of plants to infection with viruses or other pathogens appear to be involved in at least three types of defense reactions: (a) a direct attack of the invading pathogen (e.g. hydrolytic enzymes), (b) a localization of the pathogen at the site of infection (e.g. enzymes involved in lignification), and (c) an adaptation of the host metabolism to the stress condition (e.g. superoxide-dismutase). The genes encoding these inducible proteins possess interesting features, each pointing to future areas of productive research. Constitutive expression of these genes in genetically engineered plants may shed light on the role of the corresponding proteins in defense reactions and the function of their tissue-specific expression in mature or senescing plants. This insight could be used to increase the natural resistance of economically important crops.

A study of promoter regions will reveal the similarities and differences in the cis-acting elements involved in the induction of these genes by infection, elicitors, ethylene, other plant hormones, wounding, abiotic inducers, etc. Identification of the transacting factors binding to these cis-acting elements will help unravel the diversity of the induction pathways and may permit the linkage of these intracellular pathways with the intercellular signal molecules responsible for systemic induction of defense genes. Finally, the expression of mutated or chimeric proteins in plants will permit an analysis of polypeptide sequences that determine whether the protein is targeted to the vacuoles or is destined to follow the intercellular secretion pathway. Thus, studies on plant defense genes will also contribute to our general understanding of the molecular biology of plant cells.

Literature Cited

1. Ahl, P., Gianazzi, S. 1982. b-protein as a constitutive component in highly (TMV) resistant interspecific hybrids of Nicotiana glutinosa x Nicotiana debneyi. Plant Sci. Lett. 26:173–81
2. Antoniw, J. F., Ritter, C. E., Pierpoint, W. S., Van Loon, L. C. 1980. Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV. J. Gen. Virol. 47:79–87
3. Asselin, A., Grenier, J., Cote, F. 1985. Light-influenced extra cellular accumulation of b (pathogenesis-related) proteins in Nicotiana green tissue induced by various chemicals or prolonged flotation on water. Can. J. Bot. 63:1276–83
4. Bohlmann, H., Clausen, S., Behnke, S., Giese, H., Hiller, C., et al. 1988. Leaf-specific thionins of barley—a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defence mechanism of plants. EMBO J. 7:1559–65
5. Bol, J. F. 1988. Structure and expression of plant genes encoding pathogenesis-related proteins. In Plant Gene Research; Temporal and Spatial Regulation of Plant Genes, ed. D. P. S. Verma,

R. B. Goldberg, pp. 201–21. Wien/New York: Springer-Verlag
6. Bol, J. F., Linthorst, H. J. M., Meuwissen, R. L. J. 1990. Molecular responses of tobacco to virus infection. In Biochemistry and Molecular Biology of Plant Pathogen Interactions, ed. C. J. Smith. Oxford: Oxford Univ. Press. In press
7. Bol, J. F., Van de Rhee, M. D., Van Kan, J. A. L., González Jaén, M. T., Linthorst, H. J. M. 1989. Characterization of two virus-inducible plant promoters. See Ref. 74a, pp. 169–74
8. Bol, J. F., Van Kan, J. A. L. 1988. The synthesis and possible functions of virus-induced proteins in plants. Microbiol. Sci. 5:47–52
9. Bol, J. F., Van Kan, J. A. L., Cornelissen, B. J. C. 1988. Plant defense genes induced by virus infection. In Molecular Biology of Plant-Pathogen Interactions, ed. B. Staskawicz, P. Ahlquist, O. Yoder, pp. 247–54. New York: Liss
10. Bol, J. F., Van Rossum, C. M. A., Cornelissen, B. J. C., Linthorst, H. J. M. 1990. Induction of host genes by the hypersensitive response of tobacco to virus infection. In Viral Genes and Plant Pathogenesis, ed. J. G. Shaw, T. P. Pirone. Wien/New York: Springer-Verlag. In press
11. Boller, T. 1985. Induction of hydrolases as a defense reaction against pathogens. In Cellular and Molecular Biology of Plant Stress, ed. J. L. Key, T. Kosuge, pp. 247–62. New York: Liss
12. Boller, T. 1988. Ethylene and the regulation of antifungal hydrolases in plants. In Oxford Surveys of Plant Molecular and Cell Biology, ed. B. J. Miflin, 5:145–75. Oxford: Oxford Univ. Press
13. Boller, T., Metraux, J. P. 1988. Extracellular localization of chitinase in cucumber. Physiol. Mol. Plant Pathol. 33:11–16
14. Boller, T., Vögeli, U. 1984. Vacuolar localization of ethylene-induced chitinase in bean leaves. Plant Physiol. 74:442–44
15. Bowler, C., Alliotte, T., De Loose, M., Van Montagu, M., Inzé, D. 1989. The induction of manganese superoxide dismutase in response to stress in Nicotiana plubaginifolia. EMBO J. 8:31–38
16. Broglie, K. E., Biddle, P., Cressman, R., Broglie, R. 1989. Functional analysis of DNA sequences responsible for ethylene regulation of a bean chitinase gene in transgenic tobacco. Plant Cell 1:599–607
17. Broglie, K. E., Gaynor, J. J., Broglie, R. M. 1986. Ethylene-regulated gene

expression: molecular cloning of the genes encoding an endochitinase from Phaseolus vulgaris. Proc. Natl. Acad. Sci. USA 83:6820–24
18. Carr, J. P., Beachy, R. N., Klessig, D. F. 1989. Are the PR1 proteins of tobacco involved in genetically engineered resistance to TMV? Virology 169:470–73
19. Carr, J. P., Dixon, D. C., Nikolau, B. J., Voelkerding, K. V., Klessig, D. F. 1987. Synthesis and localization of pathogenesis-related proteins in tobacco. Mol. Cell. Biol. 7:1580–83
20. Collinge, D. B., Slusarenko, A. J. 1987. Plant gene expression in response to pathogens. Plant Mol. Biol. 9:389–410
21. Condit, C. M., Meagher, R. B. 1987. Expression of a gene encoding a glycine-rich protein in petunia. Mol. Cell. Biol. 7:4273–79
22. Cooper, J. B., Chen, J. A., Van Holst, G-J., Varner, J. E. 1987. Hydroxyproline-rich glycoproteins of plant cell walls. Trends Biol. Sci. 12:24–27
23. Cornelissen, B. J. C., Hooft van Huisduijnen, R. A. M., Bol, J. F. 1986. A tobacco mosaic virus-induced tobacco protein is homologous to the sweet-tasting protein thaumatin. Nature 321:531–32
24. Cornelissen, B. J. C., Hooft van Huisduijnen, R. A. M., Van Loon, L. C., Bol, J. F. 1986. Molecular characterization of messenger RNAs for "pathogenesis-related" proteins 1a, 1b and 1c, induced by TMV infection of tobacco. EMBO J. 5:37–40
25. Cornelissen, B. J. C., Horowitz, J., Van Kan, J. A. L., Goldberg, R. B., Bol, J. F. 1987. Structure of tobacco genes encoding pathogenesis-related proteins from the PR-1 group. Nucleic Acids Res. 15:6799–11
26. Cramer, C. L., Edwards, K., Dron, M., Liang, X., Dildine, S., et al. 1989. Phenylalanine ammonia-lyase gene organization and structure. Plant Mol. Biol. 12:367–83
27. Cutt, J. R., Dixon, D. C., Carr, J. P., Klessig, D. F. 1988. Isolation and nucleotide sequence of cDNA clones for the pathogenesis-related proteins PR1a, PR1b and PR1c of Nicotiana tabacum cv. Xanthi nc induced by TMV infection. Nucleic Acids Res. 16:9661
28. De Looze, M., Alliotte, T., Gheysen, G., Genetello, C., Gielen, J., et al. 1989. Primary structure of a hormonally regulated β-glucanase of Nicotiana plumbaginifolia. Gene 70:13–23
29. Dron, M., Clouse, S. D., Dixon, R. A.,

30. Lawton, M. A., Lamb, C. J. 1988. Glutathione and fungal elicitor regulation of a plant defense gene promoter in electroporated protoplasts. Proc. Natl. Acad. Sci. USA 85:6738–42

31. Ecker, J. R., Davis, R. W. 1987. Plant defense genes are regulated by ethylene. Proc. Natl. Acad. Sci. USA 84:5202–6

32. Feinbaum, R. L., Ausubel, F. M. 1988. Transcriptional regulation of the Arabidopsis thaliana chalcone synthase gene. Mol. Cell. Biol. 8:1985–92

33. Felix, G., Meins, F. 1985. Purification, immunoassay and characterization of an abundant, cytokinin-regulated polypeptide in cultured tobacco tissues. Evidence the protein is a β-1,3-glucanase. Planta 164:423–28

34. Felix, G., Meins, F. 1986. Developmental and hormonal regulation of β-1,3-glucanase in tobacco. Planta 167:206–11

35. Fincher, G. B., Lock, P. A., Morgan, M. M., Lingelbach, K., Wettenhall, R. E. H., et al 1986. Primary structure of the (1,3-1,4)-β-D-glucan 4-glucohydrolase from barley aleurone. Proc. Natl. Acad. Sci. USA 83:2081–85

36. Fraser, R. S. S., Clay, C. M. 1983. Pathogenesis-related proteins and acquired systemic resistance: causal relationship or separate effects? Neth. J. Plant Pathol. 89:283–92

37. Fritig, B., Kauffmann, S., Dumas, B., Geoffroy, P., Kopp, M., et al 1987. Mechanism of the hypersensitive reaction of plants. In Plant Resistance to Viruses, ed. D. Evered, S. Harnett. Ciba Found. Symp. 133:92–108 Chichester: Wiley & Sons

38. Fritig, B., Rouster, J., Kauffmann, S., Stinzi, A., Geoffroy, P., et al 1989. See Ref. 74a, pp. 161–68

39. Gaynor, J. J. 1988. Primary structure of an endochitinase mRNA from Solanum tuberosum. Nucleic Acids Res. 16:5210

40. Gianinazzi, S. 1983. Genetic and molecular aspects of resistance induced by infection or chemicals. In Plant Microbe Interactions; Molecular and Genetic Perspectives, ed. E. W. Nester, T. Kosuge, 1:321–42. New York: Macmillan

41. Gianinazzi, S., Ahl, P. 1983. The genetic and molecular basis of b-proteins in the genus Nicotiana. Neth. J. Plant Pathol. 89:275–81

42. Gianinazzi, S., Martin, C., Vallee, J. C. 1970. Hypersensibilité aux virus, temperatures et protéines solubles chez le Nicotiana Xanthi nc. Apparition de nouvelles macromolecules lors de la répression de la synthèse virale. C. R. Acad. Sci. Paris D 270:2383–86

43. Gianinazzi, S., Pratt, H. M., Shewry, P. R., Miflin, B. J. 1977. Partial purification and preliminary characterization of soluble proteins specific to virus infected tobacco plants. J. Gen. Virol. 34:345–51

44. Gómez, J., Sánchez-Martínez, D., Stiefel, V., Rigau, J., Puigdomènech, P., et al 1988. A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. Nature 334:262–64

45. Hedrick, S. A., Bell, J. N., Boller, T., Lamb, C. J. 1988. Chitinase cDNA cloning and mRNA induction by fungal elicitor, wounding and infection. Plant Physiol. 86:182–86

46. Herrmann, A., Schulz, W., Hahlbrock, K. 1988. Two alleles of the single-copy chalcone synthase gene in parsley differ by a transposon-like element. Mol. Gen. Genet. 212:93–98

47. Hilder, V. A., Gatehouse, A. M. R., Sheerman, S. E., Barker, R. F., Boulter, D. 1987. A novel mechanism of insect resistance engineered into tobacco. Nature 300:160–63

48. Høj, P. B., Hartman, D., Morrice, N. A., Doan, D. N. P., Fincher, G. B. 1989. Purification of (1,3)-β-glucan endohydrolase isozyme II from germinated barley and determination of its primary structure from a cDNA clone. Plant Mol. Biol. 13:31–42

49. Hooft van Huijsduijnen, R. A. M., Alblas, S. W., De Rijk, R. H., Bol, J. F. 1986. Induction by salicylic acid of pathogenesis-related proteins and resistance to alfalfa mosaic virus infection in various plant species. J. Gen. Virol. 67:2135–43

50. Hooft van Huijsduijnen, R. A. M., Kauffmann, S., Brederode, F. Th., Cornelissen, B. J. C., Legrand, M., et al 1987. Homology between chitinases that are induced by TMV infection of tobacco. Plant. Mol. Biol. 9:411–20

51. Hooft van Huijsduijnen, R. A. M., Van Loon, L. C., Bol, J. F. 1986. cDNA cloning of six mRNAs induced by TMV infection of tobacco and a characterization of their translation products. EMBO J. 5:2057–61

52. Hosokawa, D., Ohashi, Y. 1988. Immuno-chemical localization of pathogenesis-related proteins secreted into the intercellular spaces of salicylate-treated tobacco leaves. Plant Cell Physiol. 29:88–128

53. Jamet, E., Fritig, B. 1986. Purification and characterization of 8 of the pathogenesis-related proteins in tobacco leaves reacting hypersensitively to tobacco mosaic virus. Plant Mol. Biol. 6:69–80

54. Kaldenhoff, R., Richter, G. 1989. Sequence of cDNA for a novel light-induced glycin-rich protein. Nucleic Acids Res. 17:2853

55. Kauffmann, S. 1988. Les protéines PR (pathogenesis-related) du tabac: des protéines impliquées dans les réactions de défence aux agents phytopathogènes. Isolement, propriété sérologiques et activités biologiques. PhD thesis. Univ. Louis Pasteur, Strasbourg. 85 pp.

56. Kauffmann, S., Legrand, M., Geoffroy, P., Fritig, B. 1987. Biological function of "pathogenesis-related" proteins. Four PR-proteins of tobacco have 1,3-β-glucanase activity. EMBO J. 6:3209–12

57. Keller, B., Sauer, N., Lamb, C. J. 1988. Glycine-rich cell wall proteins in bean: gene structure and association of the protein with the vascular system. EMBO J. 7:3625–33

58. Keller, B., Schmid, J., Lamb, C. J. 1989. Vascular expression of a bean cell wall glycine-rich protein-β-glucuronidase gene fusion in transgenic tobacco. EMBO J. 8:1309–14

59. King, G. J., Turner, V. A., Hussey, V. E., Wurtele, E. S., Lee, S. M. 1988. Isolation and characterization of a tomato cDNA clone which codes for a salt-induced protein. Plant Mol. Biol. 10:401–12

60. Kombrink, E., Schröder, M., Hahlbrock, K. 1988. Several "pathogenesis-related" proteins in potato are 1,3-β-glucanases and chitinases. Proc. Natl. Acad. Sci. USA 85:782–86

61. Lagrimini, L. M., Burkhart, W., Moyer, M., Rothstein, S. 1987. Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: molecular analysis and tissue-specific expression. Proc. Natl. Acad. Sci. USA 84:7542–46

62. Lagrimini, L. M., Rothstein, S. 1987. Tissue specificity of tobacco peroxidase isozymes and their induction by wounding and tobacco mosaic virus infection. Plant Physiol. 84:438–42

63. Lawton, M. A., Clouse, S. D., Lamb, C. J. 1990. Gluthathione-elicited changes in chromatin configuration in the promoter of the plant defense gene chalcone synthase. EMBO J. In press

64. Lawton, M. A., Lamb, C. J. 1987. Transcriptional activation of plant defense genes by fungal elicitor, wounding and infection. Mol. Cell. Biol. 7:335–41

65. Leach, J. E., Sherwood, J., Fulton, R. W., Sequeira, L. 1983. Comparison of soluble proteins associated with disease resistance induced by bacterial lipopolysaccharide and by viral necrosis. Physiol. Plant. Pathol. 23:377–85

66. Legrand, M., Kauffmann, S., Geoffroy, P., Fritig, B. 1987. Biological function of "pathogenesis-related" protein: four tobacco PR-proteins are chitinases. Proc. Natl. Acad. Sci. USA 84:6750–54

67. Linthorst, H. J. M., Cornelissen, B. J. C., Van Kan, J. A. L., Van de Rhee, M., Meuwissen, R. L. J., et al 1990. Induction of plant genes by compatible an incompatible virus-plant interactions. In Recognition and Response in Plant-Virus Interactions, ed. R. S. S. Fraser, pp. 361–73. Wien/New York: Springer-Verlag

68. Linthorst, H. J. M., Meuwissen, R. L. J., Kauffmann, S., Bol, J. F. 1989. Constitutive expression of pathogenesis-related proteins PR-1, GRP and PR-S in tobacco has no effect on virus infection. Plant Cell 1:285–91

69. Lois, R., Dietrich, A., Hahlbrock, K., Schulz, W. 1989. A phenylalanine amonia-lyase gene from parsley: structure, regulation and identification of elicitor and light-responsive cis-acting elements. EMBO J. 8:1641–48

70. Lotan, T., Ori, N., Fluhr, R. 1988. Pathogenesis Related Proteins Are Developmentally Regulated in Tobacco Flowers. Presented at 2nd Int. Congr. Plant Mol. Biol., Jerusalem

71. Lotan, T., Ori, N., Fluhr, R. 1989. Pathogeneis related proteins are developmentally regulated in tobacco flowers. Plant Cell 1:881–87

72. Lozoya, E., Hoffmann, H., Douglas, C., Schulz, W., Scheel, D. et al 1988. Primary structures and catalytic properties of isozymes encoded by the two 4-coumarate:CoA ligase genes in parsley. Eur. J. Biochem. 176:661–67

73. Lucas, J., Henriquez, A. C., Lottspeich, F., Henschen, A., Sanger, H. L. 1985. Amino acid sequence of the "pathogenesis-related" leaf protein p14 from viroid-infected tomato reveals a new type of structurally unfamiliar proteins. EMBO J. 4:2745–49

74. Lucas, J., Henschen, A., Lottspeich, F., Voegeli, U., Boller, T. 1985. Amino-terminal sequence of ethylene-induced bean leaf chitinase reveals similarities to sugar-binding domains of wheat germ agglutinin. *FEBS Lett.* 193:208–10

74a. Lugtenberg, B. J. J., ed. *Signal Molecules in Plants and Plant-Microbe Interactions, NATO ASI Ser.*, H: Cell Biol. Vol. 36. Wien: Springer-Verlag

75. Maiss, E., Poehling, H. M. 1983. Resistance against plant viruses induced by culture filtrates of the fungus *Stachybotrys chartarum. Neth. J. Plant Pathol.* 89:323

76. Matsuoka, M., Yamamoto, N., Kano-Murakami, Y., Tanaka, Y., Ozeki, Y., et al. 1987. Classification and structural comparison of full-length cDNAs for pathogenesis-related proteins. *Plant Physiol.* 85:942–46

77. Mauch, F., Mauch-Mani, B., Boller, T. 1988. Antifungal hydrolases in pea tissue II. Inhibition of fungal growth by combinations of chitinase and β-1,3-glucanase. *Plant Physiol.* 88:936–42

78. Mauch, F., Staehelin, L. A. 1989. Functional implications of the subcellular localization of ethylene-induced chitinase and β-1,3-glucanase in bean leaves. *Plant Cell* 1:447–57

79. Memelink, J. 1988. *Altered gene expression in T-DNA transformed tobacco tissues.* PhD thesis. Leiden Univ., Leiden. 126 pp.

80. Memelink, J., Linthorst, H. J. M., Schiperoort, R. A., Hoge, J. H. C. 1990. Tobacco genes encoding acidic and basic pathogenesis-related proteins display different expression patterns. *Plant Mol. Biol.* 14:119–26

81. Metraux, J. P., Burkhart, W., Moyer, M., Dincher, S., Middlesteadt, W., et al. 1989. Isolation of a complementary DNA encoding a chitinase with structural homology to a bifunctional lysozyme/chitinase. *Proc. Natl. Acad. Sci. USA* 86:896–900

82. Mozes, R., Anignus, Y., Sela, I., Harpaz, I. 1978. The chemical nature of an antiviral factor (AVF) from virus-infected plants. *J. Gen. Virol.* 38:241–49

83. Mundy, J., Chua, N-H. 1988. Abscisic acid and water-stress induce the expression of a novel rice gene. *EMBO J.* 7:2279–86

84. Nasser, W., de Tapia, M., Kauffmann, S., Montasser-Kouhsari, S., Burkhard, G. 1988. Identification and characterization of maize pathogenesis-related pro-

teins. Four maize PR proteins are chitinases. *Plant Mol. Biol.* 11:529–38

85. Nassuth, A., Sanger, H. L. 1986. Immunological relationships between "pathogenesis-related" leaf proteins from tomato, tobacco and cowpea. *Virus Res.* 4:229–42

86. Ohashi, Y., Matsuoka, M. 1987. Localization of pathogenesis-related proteins in the epidermis and intercellular spaces of tobacco leaves after their induction by potassium salicylate or tobacco mosaic virus infection. *Plant Cell Physiol.* 28:1227–35

87. Ohshima, M., Matsuoka, M., Yamamoto, N., Tanaka, Y., Kano-Murakami, Y., et al 1987. Nucleotide sequence of the PR-1 gene of Nicotiana tabacum. *FEBS Lett.* 225:243–46

88. Parent, J. G., Asselin, A. 1984. Detection of pathogenesis-related proteins (PR or *b*) and of other proteins in the intercellular fluid of hypersensitive plants infected with tobacco mosaic virus. *Can. J. Bot.* 62:564–69

89. Payne, G., Middlesteadt, W., Desai, N., Williams, S., Dincher, S., et al. 1989. Isolation and sequence of a genomic clone encoding the basic form of pathogenesis-related protein 1 from *Nicotiana tabacum. Plant Mol. Biol.* 12:595–96

90. Payne, G., Middlesteadt, W., Williams, S., Desai, N., Parks, T. D., et al. 1988. Isolation and nucleotide sequence of a novel cDNA clone encoding the major form of pathogenesis-related protein R. *Plant Mol. Biol.* 11:223–24

91. Payne, G., Parks, T. D., Burkhart, W., Dincher, S., Ahl, P., et al. 1988. Isolation of the genomic clone for pathogenesis-related protein *b* from *Nicotiana tabacum* cv. Xanthi *nc. Plant Mol. Biol.* 11:89–94

92. Pfitzner, U. M., Goodman, H. M. 1987. Isolation and characterization of cDNA clones encoding pathogenesis-related proteins from tobacco mosaic virus infected tobacco plants., *Nucleic Acids Res.* 15:4449–65

93. Pfitzner, U. M., Pfitzner, A. J. P., Goodman, H. M. 1988. DNA sequence analysis of a PR-1a gene from tobacco: molecular relationship of heat shock and pathogen responses in plants. *Mol. Gen. Genet.* 211:29095

94. Pierpoint, W. S. 1986. The pathogenesis-related proteins of tobacco leaves. *Phytochemistry* 25:1595–601

95. Pierpoint, W. S., Tatham, A. S., Pappin, D. J. C. 1987. Identification of the virus-induced protein of tobacco leaves

96. Raikhel, N. V., Wilkins, T. A., Mansfield, M. A. 1988. Molecular basis of post-translational modification of wheat germ agglutinin. *Curr. Top. Plant Biochem. Physiol.* 7:83–89

97. Richardson, M., Valdes-Rodriguez, S., Blanco-Labra, A. 1987. A possible function for thaumatin and a TMV-induced protein suggested by homology to a maize inhibitor. *Nature* 327:432–34

98. Roberts, E., Kolattukudy, P. E. 1989. Molecular cloning, nucleotide sequence and abscisic acid induction of a suberization-associated highly anionic peroxidase. *Mol. Gen. Genet.* 217:223–32

99. Roberts, E., Kutchan, T., Kolattukudy, P. E. 1988. Cloning sequencing of cDNA for a highly anionic peroxidase from potato and the induction of its mRNA in suberizing potato tubers and tomato fruits. *Plant Mol. Biol.* 11:15–26

100. Rodrigo, I., Vera, P., Conejero, V. 1990. Degradation of tomato pathogenesis-related (PR) proteins by an endogenous 37 kDa aspartyl endoprotease. *Eur. J. Biochem.* In press

101. Ryan, C. A. 1987. Oligosaccharide signalling in plants. *Annu. Rev. Cell Biol.* 3:295–317

102. Schlumbaum, A., Mauch, F., Vögeli, U., Boller, T. 1986. Plant chitinases are potent inhibitors of fungal growth. *Nature* 324:365–67

103. Shinshi, H., Mohnen, D., Meins, F. 1987. Regulation of a plant pathogenesis-related enzyme: inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin. *Proc. Natl. Acad. Sci. USA* 84:89–93

104. Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J., et al. 1988. Evidence for N- and C-terminal processing of a plant defense-related enzyme primary structure of tobacco prepro-β-1,3-glucanase. *Proc. Natl. Acad. Sci. USA* 85:5541–45

105. Singh, N. K., Bracker, C. A., Hasegawa, P. M., Handa, A. K., Buckel, S., et al. 1987. Characterization of osmotin. A thaumatin-like protein associated with osmotic adaptation in plant cells. *Plant Physiol.* 85:529–36

106. Somssich, I. E., Bollmann, J., Hahlbrock, K., Kombrink, E., Schulz, W. 1989. Differential early activation of defense-related genes in elicitor-treated parsley cells. *Plant Mol. Biol.* 12:227–34

107. Somssich, I. E., Schmelzer, E., Kawalleck, P., Hahlbrock, K. 1988. Gene structure and in situ transcript localization of pathogenesis-related protein 1 in parsley. *Mol. Gen. Genet.* 213:93–98

108. Spiegel, S., Gera, A., Salomon, P., Ahl, P., Harlap, S., et al. 1989. Recovery of an inhibitor of virus replication from the intercellular fluid , of hypersensitive tobacco infected with tobacco mosaic virus and from un-infected induced-resistant tissue. *Phytopathology* 79:258–62

109. Stanford, A., Bevan, M., Northcote, D. 1989. Differential expression within a family of novel wound-induced genes in potato. *Mol. Gen. Genet.* 215:200–208

110. Swegle, M., Huang, J-K., Lee, G., Muthukrishnan, S. 1989. Identification of an endochitinase cDNA clone from barley aleurone cells. *Plant Mol. Biol.* 12:403–12

111. Takahashi, H., Shimamoto, K., Ehara, Y. 1989. Cauliflower mosaic virus gene VI causes growth suppression, development of necrotic spots and expression of defence-related genes in transgenic tobacco plants. *Mol. Gen. Genet.* 216: 188–94

112. Van den Bulcke, M., Bauw, G., Castresana, C., Van Montagu, M., Vandekerckhove, J. 1989. Characterization of vacuolar and extracellular β-(1,3)-glucanases of tobacco: evidence for a strictly compartmentalized plant defense system. *Proc. Natl. Acad. Sci. USA* 86:2673–77

113. Van Kan, J. A. L., Cornelissen, B. J. C., Bol, J. F. 1988. A virus-inducible tobacco gene encoding a glycine-rich protein shares putative regulatory elements with the ribulose bisphosphate carboxylase small subunit gene. *Mol. Plant-Microbe Interact.* 1:107–12

114. Van Kan, J. A. L., Van de Rhee, M. D., Zuidema, D., Cornelissen, B. J. C., Bol, J. F. 1989. Structure of tobacco genes encoding thaumatin-like proteins. *Plant Mol. Biol.* 12:153–55

115. Van Loon, L. C. 1976. Specific soluble leaf proteins in virus-infected tobacco plants are not normal constituents. *J. Gen. Virol.* 309:375–79

116. Van Loon, L. C. 1977. Induction by 2-chloroethylphosphonic acid of viral-like lesions, associated proteins, and systemic resistance in tobacco. *Virology* 80:417–20

117. Van Loon, L. C. 1982. Regulation of changes in proteins and enzymes associated with active defence against virus infection. In *Active Defense Mechanisms*

LONG-DISTANCE DISPERSION OF RUST PATHOGENS

S. Nagarajan

Indian Council of Agricultural Research, Krishi Bhavan, New Delhi 110 001, India

D. V. Singh

Division of Mycology and Plant Pathology, Indian Agricultural Research Institute, New Delhi 110 012, India

KEY WORDS: trajectory, remote sensing, climatic factors, *Puccinia* path.

INTRODUCTION

Migration and dispersal are the two main evolutionary strategies adopted by living organisms to colonize new sites. The migration of birds and mammals in search of new sites and sources of food occurs in an orderly, coordinated way with minimum wastage of progenies. Plant pathogens, on the other hand, produce enormous numbers of spores that are passively transported, scattered in all directions, and finally land on nontarget sites in uncongenial environments as well as on congenial hosts. Although the dispersion of rust uredio-spores may seem a wasteful exercise, this random profligacy is compensated for by the enormous numbers of spores produced per unit area. The purpose of this chapter is to describe some of the evidence and routes by which long-distance dispersal (LDD) of plant pathogens is known to occur in various parts of the world. The chapter also contains a description of new tools by which LDD can be studied and made into a predictive tool for advisory services.

DEFINING THE TERMS

Source area is the suspected or proven place or warehouse from which inoculum is dispersed over both short and long distances. *Target* is a dense

139

in Plants, ed. R. K. S. Wood, pp. 247–73. New York: Plenum

118. Van Loon, L. C. 1985. Pathogenesis-related proteins. *Plant Mol. Biol.* 4:111–16

119. Van Loon, L. C. 1990. Stress proteins in infected plants. In *Plant-Microbe Interactions; Molecular and Genetic Perspectives*, ed. T. Kosuge, E. W. Nester, New York: Macmillan. In press

120. Van Loon, L. C., Gerritsen, Y. A. M., Ritter, C. E. 1987. Identification, purification and characterization of pathogenesis-related proteins from virus-infected Samsun NN tobacco leaves. *Plant Mol. Biol.* 9:593–609

121. Van Loon, L. C., Van Kammen, A. 1970. Polyacrylamide disc electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var "Samsun" and "Samsun NN". II. Changes in protein constitution after infection with tobacco mosaic virus. *Virology* 40:199–211

122. Van Tunen, A. J., Hartman, S. A., Mur, A. Mol, J. N. M. 1989. Regulation of chalcone flavanone isomerase (CHI) gene expression in *Petunia hybrida*: the use of alternative promoters in corolla, anthers and pollen. *Plant Mol. Biol.* 12:539–51

123. Vera, P., Conejero, V. 1988. Pathogenesis-related proteins of tomato. P-69 as an alkaline endoproteinase. *Plant Physiol.* 87:58–63

124. Vera, P., Hernández-Yago, J., Conejero, V. 1989. "Pathogenesis-related" P1(p14) protein. Vacuolar and apoplastic localization in leaf tissue from tomato plants infected with citrus exocortis viroid; *in vitro* synthesis and processing. *J. Gen. Virol.* 70:1933–42

125. Verner, K., Schatz, G. 1988. Protein translocation across membranes. *Science* 241:1307–13

126. Voelker, T. A., Herman, E. M., Chrispeels, M. J. 1989. *In vitro* mutated phytohemagglutinin genes expressed in tobacco seeds: role of glycans in protein targeting and stability. *Plant Cell* 1:95–104

127. Vögeli, U., Meins, F., Boller, T. 1988. Co-ordinated regulation of chitinase and β-1,3-glucanase in bean leaves. *Planta* 174:364–72

128. Walter, M. H., Grima-Pettenati, J., Grand, C., Boudet, A. M., Lamb, C. J. 1988. Cinamyl-alcohol dehydrogenase, a molecular marker specific for lignin synthesis: cDNA cloning and mRNA induction by fungal elicitor. *Proc. Natl. Acad. Sci. USA* 85:5546–50

129. Watson, M. E. E. 1984. Compilation of published signal sequences. *Nucleic Acids Res.* 12:5145–64

130. White, R. F., Rybicki, E. P., Von Wechmar, M. B., Dekker, J. L., Antoniw, J. F. 1987. Detection of PR-1 type proteins in *Amaranthaceae, Chenopodiaceae, Gramineae* and *Solanaceae* by immunoelectroblotting. *J. Gen. Virol.* 68:2043–48

131. Wingender, R., Röhrig, H., Höricke, C., Wing, D., Schell, J. 1989. Differential regulation of soybean chalcone synthase genes in plant defence, symbiosis and upon environmental stimuli. *Mol. Gen. Genet.* 218:315–22

Sorting of Phaseolin to the Vacuole Is Saturable and Requires a Short C-Terminal Peptide

Lorenzo Frigerio,[a,b] Maddalena de Virgilio,[a,1] Alessandra Prada,[a] Franco Faoro,[c] and Alessandro Vitale[a,2]

[a]Istituto Biosintesi Vegetali, Consiglio Nazionale delle Ricerche, via Bassini 15, 20133 Milan, Italy
[b]Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom
[c]Centro Miglioramento Sanitario Colture Agrarie, Consiglio Nazionale delle Ricerche, via Celoria 2, 20133 Milan, Italy

Phaseolin, one of the major legume proteins for human nutrition, is a trimeric glycoprotein of the 7S class that accumulates in the protein storage vacuoles of common bean. Phaseolin is cotranslationally introduced into the lumen of the endoplasmic reticulum; from there, it is transported through the Golgi complex to the storage vacuoles. Phaseolin is also transported to the vacuole in vegetative tissues of transgenic plants. By transient and permanent expression in tobacco leaf cells, we show here that vacuolar sorting of phaseolin is saturable and that saturation leads to Golgi-mediated secretion from the cell. A mutated phaseolin, in which the four C-terminal residues (Ala, Phe, Val, and Tyr) were deleted, efficiently formed trimers but was secreted entirely outside of the cells in transgenic tobacco leaves, indicating that the deleted sequence contains information necessary for interactions with the saturable vacuolar sorting machinery. In the apoplast, the secreted phaseolin remained intact; this is similar to what occurs to wild-type phaseolin in bean storage vacuoles, whereas in vegetative vacuoles of transgenic plants, the storage protein is fragmented.

INTRODUCTION

Phaseolin is the major storage protein of common bean. Phaseolin is a member of the 7S vicilin class and one of the most important legume proteins for human nutrition; a number of efforts have been made to improve its nutritional value (Hoffman et al., 1988; Dyer et al., 1995). The structure, genetic makeup, cotranslational and post-translational modifications, and intracellular transport of phaseolin have been elucidated largely by numerous investigators (Bollini et al., 1982; Slightom et al., 1985; Sturm et al., 1987; Lawrence et al., 1994), but the mechanisms that allow correct intracellular targeting of phaseolin and the other 7S storage proteins have not been fully characterized.

Phaseolin is a homotrimeric soluble protein that accumulates in the protein storage vacuoles of cotyledonary cells. Its synthesis, maturation, and intracellular targeting are mediated by the secretory pathway, which delivers proteins into the endoplasmic reticulum (ER) and from there to the cell surface or the vacuoles (Okita and Rogers, 1996). The Golgi complex as well as other intermediate compartments mediate this traffic.

Protein constructs that have a transient signal peptide for cotranslational insertion into the ER, but no other specific sorting signal, are secreted from plant cells (Denecke et al., 1990; Hunt and Chrispeels, 1991). Soluble proteins destined for the different vacuoles are sorted from the proteins destined for the apoplast, probably at the exit of the Golgi complex (Ahmed et al., 1997; Paris et al., 1997). Sorting occurs because vacuolar proteins have structures, often identified as short stretches of amino acids in propeptides, that are not present in apoplastic proteins. The signals are variable, and different vacuolar sorting mechanisms must exist (Matsuoka et al., 1995; Kirsch et al., 1996). A putative integral membrane receptor that recognizes some but not all of these stretches in vitro has been identified (Kirsch et al., 1996; Ahmed et al., 1997).

By expressing different phaseolin constructs, we show here that although phaseolin synthesized in cells of transgenic tobacco leaves is exclusively vacuolar, vacuolar targeting of this protein can be saturated at high transient expression levels in the same cell types, resulting in the Golgi-mediated secretion of the protein. We also show that removal of the four C-terminal residues (Ala, Phe, Val, and Tyr) allows correct trimer formation but causes phaseolin to be entirely secreted outside of the cells in transgenic tobacco leaves and that this C-terminal sequence is most probably exposed on the surface of wild-type phaseolin trimers. Wild-type phaseolin is fragmented in vegetative vacuoles of transgenic plants (Murai et al., 1983; Sengupta-Gopalan et al., 1985; Bagga et al., 1992; Pedrazzini et al., 1997), whereas the mutated construct secreted into the

apoplast is not subjected to fragmentation. Our results are consistent with the presence of a saturable vacuolar sorting mechanism that recognizes a discrete signal located at the C terminus of phaseolin. In transgenic tobacco leaves that express wild-type phaseolin, this mechanism operates close to its saturation limit.

RESULTS

Assembly-Competent Phaseolin Is Partially Secreted from Transiently Transformed Protoplasts

This study was performed using the mutated phaseolin T343F and other mutants constructed in a T343F background. Figure 1 shows the sequence of wild-type β-phaseolin and of the mutated constructs used in this work (from amino acid 341 to the C terminus); we have not introduced any mutation in the preceding sequence. In T343F, we inactivated the second of the two N-glycosylation sites of a wild-type β-phaseolin sequence; this mutated phaseolin is therefore glycosylated with a single glycan (Pedrazzini et al., 1997). The mutation has no negative effect on the assembly and intracellular transport of phaseolin (Pedrazzini et al., 1997), and indeed in wild-type phaseolin, the second glycosylation site is a "weak" site, which is glycosylated only in some of the molecules. The advantage of using T343F rather than wild-type phaseolin resides in the fact that modifications of the single glycan by enzymes located in the Golgi complex can be used to trace the intracellular transport of the protein (Sturm et al., 1987).

While examining the intracellular transport of mutated phaseolin polypeptides, which we produced to study the relationships between trimerization and intracellular traffic, we observed that assembly-competent polypeptides, but not assembly-defective ones, could be detected in part in the incubation medium of transiently transformed tobacco protoplasts. This is shown in Figure 2A. Tobacco mesophyll protoplasts were transiently transformed with plasmid without inserts, as a control, or with plasmid carrying the se-

quence encoding T343F or T343FΔ360 (hereafter referred to as Δ360). The latter is an assembly-defective deletion mutant of phaseolin (Figure 1; Pedrazzini et al., 1997). Protoplasts were subjected to pulse–chase labeling with ^{35}S-labeled methionine and cysteine. Finally, we analyzed phaseolin by SDS-PAGE and fluorography after immunoprecipitation from cell homogenates or incubation media.

The polypeptides detected when the control plasmid without inserts was used represent nonspecific contaminants recognized by the antiserum. Of these, the major immunoprecipitable contaminant is a polypeptide of 40 kD (Figure 2A, lanes 1 and 2); the synthesis of this polypeptide must be induced largely by the stress imposed by transient transformation, because there was almost no material at 40 kD when immunoprecipitations were performed using untreated protoplasts from transgenic leaves (Figure 2B).

After a 1-hr pulse, T343F synthesized during transient expression was detectable as an abundant band of 46 kD and less abundant fragments of 20 to 25 kD (Figure 2A, lane 3). After a 5-hr chase, a relevant proportion of intact phaseolin was converted into the smaller fragments (Figure 2A, cf. lanes 3 and 4). Post-translational fragmentation is a characteristic of phaseolin expressed in heterologous plants (Bagga et al., 1992), although in bean cotyledons, the storage protein does not undergo such fragmentation. Fragmentation is a result of transport to the heterologous vacuoles (Pedrazzini et al., 1997). A relevant amount of T343F phaseolin was also recovered unfragmented from the incubation medium, where it accumulated during the chase (Figure 2A, lanes 9 and 10). After a 5-hr chase, secreted T343F represented ~25% of the amount synthesized during the pulse and almost no intracellular intact phaseolin remained, suggesting that the rest had been targeted to the vacuole. The overall amount of T343F recovered after the chase in this and similar experiments was lower than that recovered after the pulse, most probably because of full degradation of a proportion of phaseolin in the vacuole.

Assembly-defective forms of phaseolin that remain monomeric are unable to be transported along the secretory pathway in transgenic plants (Pedrazzini et al., 1997). During transient expression, the assembly-defective construct Δ360,

```
            343                360                                                                            418
            |                  |                                                                              |
wild type   NFTGFGINANNNNRNLLAGKTDNVISSIGRALDGKDVLGLTFSGSGDEVMKLINKQSGSYFVDAHHHQQEQQKGRKGAFVY
T343F       ..F.........................................................................................
Δ418        ..F.........................................................................................*
Δ360        ..F................*
```
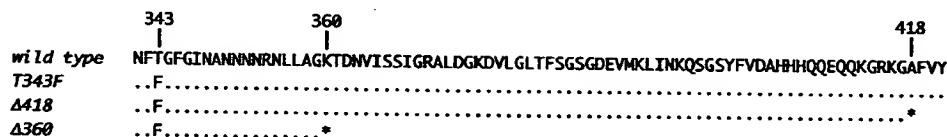
Figure 1. Comparison of the Amino Acid Sequence of Wild-Type β-Phaseolin and the Mutated Constructs Used in This Work.

The single letter code from residue 341 to the C terminus is used. The numbers on top identify the positions of the residues, starting from the first methionine of the phaseolin precursor (signal peptide included). Dots indicate identical residues. Asterisks indicate the stop codons introduced in Δ418 and Δ360. The Thr-to-Phe mutation at position 343 destroys the consensus for N-glycosylation at position 341.
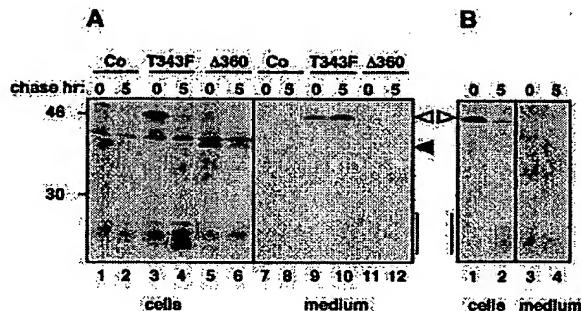
Figure 2. Transport-Competent Phaseolin Is in Part Secreted from Transiently Transfected Protoplasts but Not from Transgenic Protoplasts.

(A) Tobacco leaf protoplasts were transiently transfected with plasmid encoding the phaseolin constructs T343F or Δ360 or with a plasmid without an insert (Co). Cells were pulse-labeled with ^{35}S-methionine and ^{35}S-cysteine for 1 hr and chased for the indicated periods of time. Cells and the corresponding incubation media were then homogenized, subjected to immunoprecipitation with the anti-phaseolin antiserum, and analyzed by SDS-PAGE and fluorography. The open arrowhead indicates intact T343F; the filled arrowhead indicates Δ360; and the vertical bar indicates phaseolin fragmentation products. Numbers at left indicate molecular mass markers in kilodaltons.

(B) Protoplasts from leaves of transgenic tobacco expressing T343F were subjected to pulse–chase and analyzed as described in **(A)**. Symbols are as given in **(A)**.

which has a molecular mass of ~38 kD, was entirely recovered intracellularly, and we could not detect it in the incubation medium (Figure 2A, lanes 5 and 6, and 11 and 12). During the 5-hr chase, the protein was in part degraded: we have established that this degradation does not involve trafficking along the secretory pathway but that it is due to the not yet fully clarified mechanism of ER quality control that retains and eventually degrades defective polypeptides (Pedrazzini et al., 1997). Therefore, during transient expression, assembly-competent phaseolin is in part secreted, whereas a form of phaseolin subjected to quality control is not.

Transgenic Tobacco Plants Do Not Secrete Phaseolin

We have produced transgenic plants constitutively expressing T343F (Pedrazzini et al., 1997). As shown in Figure 2B, mesophyll protoplasts isolated from these plants do not secrete phaseolin to detectable amounts, suggesting fully efficient vacuolar sorting.

To confirm that in intact mesophyll tissue the location of T343F is exclusively vacuolar and not extracellular, we analyzed leaf tissue by immunoelectron microscopy. T343F was detected exclusively in the vacuolar lumen in large electron-

dense protein bodies (Figure 3D) that could be decorated by the anti-phaseolin antiserum (Figures 3A and 3B) but not by the preimmune serum (Figure 3C). Similar protein bodies were not found in cells of tobacco transformed with the plasmid without an insert (data not shown). No labeling above the background was detectable in the cell wall or the intercellular spaces (Figures 3A and 3B). From the results shown in Figures 2B and 3, we conclude that in transgenic leaves, T343F phaseolin is efficiently targeted to the vacuoles, where it is fragmented and forms protein bodies and is not secreted at detectable levels.

The Golgi Complex Mediates Secretion of Phaseolin from Transiently Transformed Protoplasts

One possible explanation for the different behavior of phaseolin during transient versus permanent expression is that the high expression level reached in the transiently transformed cells saturates the vacuolar sorting mechanism of phaseolin, leading to default traffic along the secretory pathway to the cell surface. In the following experiments, we tested this hypothesis.

We first determined whether secreted phaseolin travels through the Golgi complex, which is the usual route to secretion, rather than following some "nonclassical" route to the cell surface. The Golgi complex contains many glycosidases and glycosyltransferases, and one of its functions is the modification of glycans of glycoproteins. When phaseolin has only one oligosaccharide chain, its glycan undergoes processing during passage through the Golgi complex (Sturm et al., 1987). This processing makes the glycan resistant to digestion by endoglycosidase H, a fungal enzyme that removes unprocessed Asn-linked glycans from glycoproteins. Resistance to digestion by endoglycosidase H in vitro is widely used as a tool to investigate whether a protein has encountered Golgi complex–processing enzymes.

Protoplasts transiently transformed with plasmid encoding T343F were subjected to pulse–chase labeling, and phaseolin was immunoprecipitated and treated with endoglycosidase H or without the enzyme as control. At the end of the pulse, intact T343F was largely susceptible to endoglycosidase H digestion, as indicated by its slightly lower apparent molecular mass after digestion (Figure 4, cf. lanes 3 and 4). A very small proportion of phaseolin comigrated with the deglycosylated form even in the undigested control (this is visible in Figure 2A, lane 3, as well), indicating that a small proportion of polypeptides was synthesized in the unglycosylated form. After 5 hr of chase, the phaseolin polypeptides that had not yet been fragmented remained largely susceptible to digestion, indicating that most of them had not yet traveled through the Golgi complex (Figure 4, lane 6). However, the apparent molecular mass of the three major fragmentation products was not affected by endoglycosidase H (Figure 4, cf. the 20- to 25-kD fragments in lanes 5 and 6).
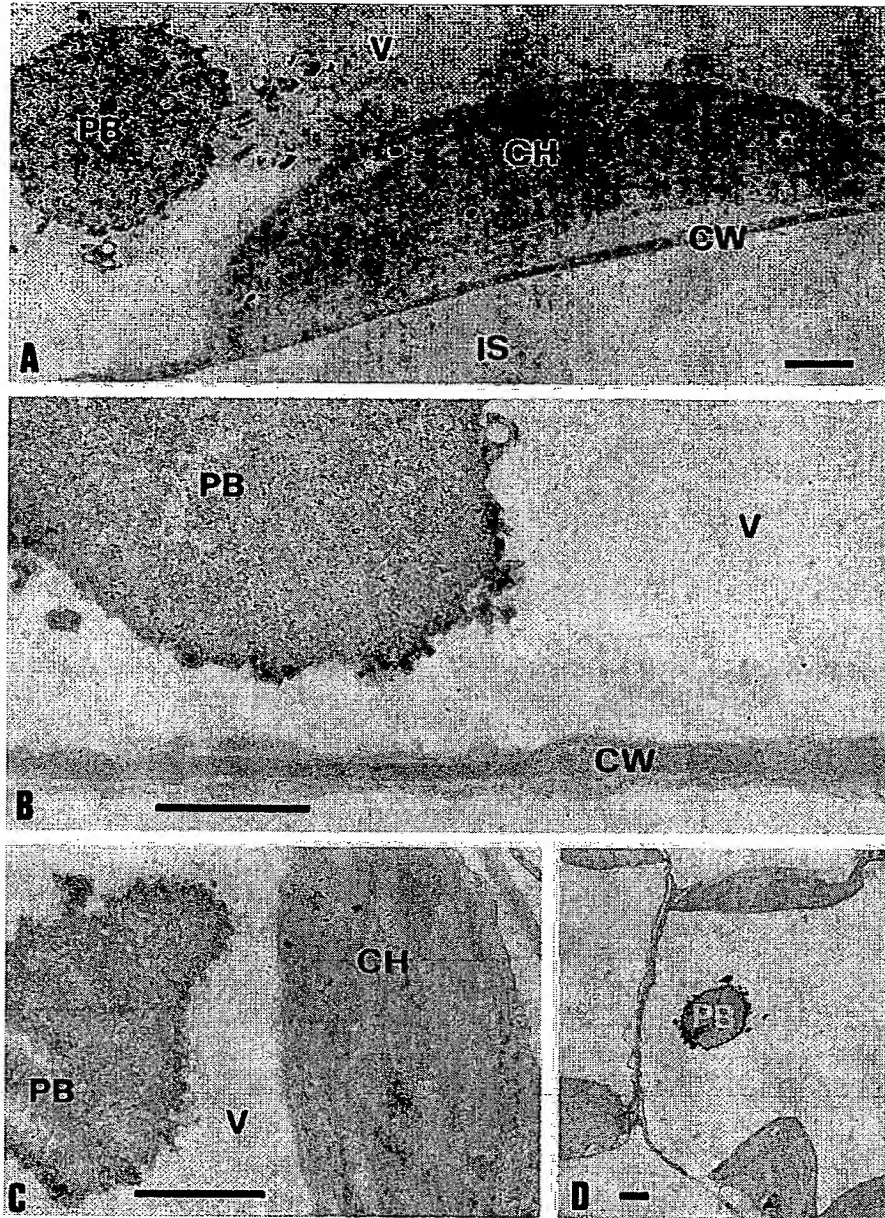
Figure 3. In Transgenic Tobacco Leaf Cells, T343F Is Located in Vacuoles.

(A) to (C) Thin sections of transgenic tobacco leaves expressing T343F were treated with anti-phaseolin antiserum ([A] and [B]) or preimmune serum (C), and the bound antibodies were visualized by using goat anti-rabbit 15-nm gold complex.

(D) A lower magnification of an untreated sample.

CH, chloroplast; CW, cell wall; IS, intercellular space; PB, protein bodies formed by phaseolin; V, vacuole. Bars = 1 μm.
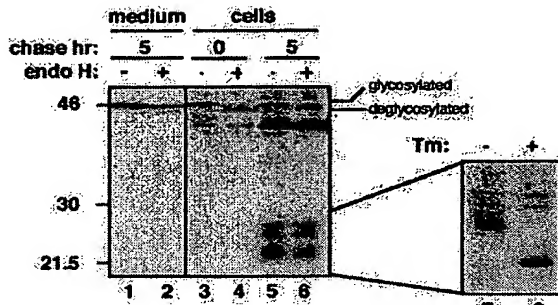
Figure 4. Both Vacuolar and Secreted T343F Acquire Endoglycosidase H Resistance.

Lanes 1 to 6 show endoglycosidase H treatment of phaseolin. Protoplasts were transiently transfected with the plasmid encoding T343F, pulse-labeled with 35S-methionine and 35S-cysteine for 1 hr, and chased for the indicated periods of time. Phaseolin was immunoselected from the cell homogenates or the incubation media and then treated with endoglycosidase H (endo H; +) or without enzyme as control (−) and analyzed by SDS-PAGE and fluorography. Lanes 5 and 6 contain material from a higher number of protoplasts than are contained in lanes 3 and 4, to allow clearer detection of the phaseolin fragmentation products. Numbers at left indicate molecular mass markers in kilodaltons. Lanes 7 and 8 show the effect of tunicamycin on the relative molecular mass of the phaseolin fragments. Protoplasts were transiently transfected with plasmid encoding T343F and were subjected to pulse–chase in the presence of tunicamycin (Tm; +) or in the absence of the inhibitor (−). Pulse was for 1 hr with 35S-methionine and 35S-cysteine and was followed by 5 hr of chase. At the end of the chase, phaseolin was immunoselected from the cell homogenates and analyzed by SDS-PAGE and fluorography. Only the portion of the gel relative to the phaseolin fragments is shown.

To determine whether any of the fragments contain the glycan, we performed radioactive labeling of protoplasts in the presence of tunicamycin, a fungal inhibitor of N-glycosylation. In the presence of the inhibitor, there was a marked decrease in the apparent molecular mass of the faster migrating phaseolin fragment, indicating that in normal conditions, this fragment is glycosylated, whereas the two slower migrating ones are not (Figure 4, cf. lanes 7 and 8). Therefore, in tobacco, as in bean cotyledons, the glycan of singly glycosylated phaseolin acquires a complex structure along the route to the vacuole. Secreted T343F was almost completely resistant to endoglycosidase H (Figure 4, cf. lanes 1 and 2). The proportion of resistant polypeptides was much higher in secreted T343F than in the intracellular T343F that was still unfragmented after the chase, ruling out the possibility that secretion of phaseolin resulted from some nonclassical "shortcut" delivery from the ER directly to the cell surface. We conclude that the Golgi complex is involved in the secretion of phaseolin.

Phaseolin Is Secreted Only When Expressed at High Levels

If secretion of phaseolin results from saturation of the vacuolar targeting mechanism, we expected a reduction in secretion when the level of expression was lowered. As a test, we transformed protoplasts with different amounts of plasmid (Figure 5). In our standard transformation protocol, we use 40 μg of plasmid to transform 10^6 protoplasts. This leads to secretion levels of T343F that vary from 20 to 50% after 5 hr of chase, using different preparations of protoplasts. Figure 5 shows that when 2 μg was used, secretion was below our limit of detection, although phaseolin was still sorted to the vacuole, as indicated by the accumulation of fragmentation products during the chase. When 20 μg of plasmid was used to transform the same preparation of protoplasts, expression of T343F was six times higher and part of the protein was secreted (Figure 5).

Therefore, the secretion of phaseolin is dose dependent, and it occurs when protoplasts are transiently transformed, but it does not occur when protoplasts are isolated from transgenic plants and are not subjected to transient transformation. This could indicate two possibilities: (1) the stress imposed on the cells by the transient transformation procedure has the side effect of altering the capacity of the vacuolar sorting mechanism; and (2) secretion is truly caused by the sudden burst of phaseolin expression, which saturates
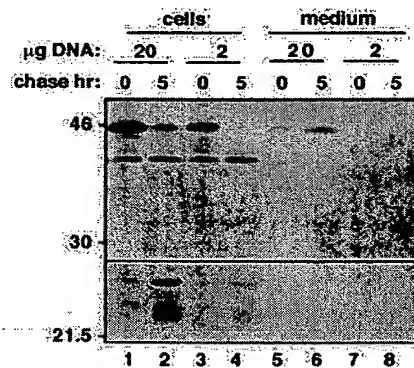
Figure 5. Secretion of T343F Is Dose Dependent.

Tobacco protoplasts were transiently transfected with 20 or 2 μg of the plasmid encoding T343F and then pulse-labeled for 1 hr with 35S-methionine and 35S-cysteine and chased for the indicated periods of time. Cell homogenates and the corresponding incubation media were then immunoprecipitated with the anti-phaseolin antiserum and analyzed by SDS-PAGE and fluorography. Numbers at left indicate molecular mass markers in kilodaltons. At bottom is a longer exposure with respect to the fluorograph at top, to allow clearer detection of the phaseolin fragmentation products (indicated by the vertical bar at right).

the normal vacuolar sorting capacity of the cells. To distinguish between these possibilities, we supertransformed protoplasts isolated from transgenic plants expressing T343F with the plasmid containing the same construct or with the plasmid without insert as a control. Transfection with the control plasmid did not change the destiny of phaseolin, whereas the overexpression of phaseolin resulted in secretion (Figure 6, lanes 7 and 8; cf. with lanes 5 and 6). The very small amount of intact phaseolin recovered in the medium after the pulse shown in lane 5 represents contamination from a small proportion of broken cells; consistently, there is no secreted phaseolin after the chase shown in lane 6.

The average overexpression of T343F as a result of the supertransformation was ~2.5-fold in the experiment shown in Figure 6. However, overexpression is probably much higher in the cells receiving the transfected DNA, because these are usually a small proportion of the total (Denecke et al., 1989). Secretion of the extra amount of phaseolin synthesized because of the transient transformation was almost quantitative. The amount of fragmentation products accumulated intracellularly after the chase was very similar in control-transformed and supertransformed protoplasts (Figure 6, cf. lanes 2 and 4). These results indicate that secretion is the result of saturation of the sorting machinery and that the mechanism that targets phaseolin to the vacuole is very close to its saturation in the transgenic plants.

Deletion of the C-Terminal Sequence Ala–Phe–Val–Tyr Causes Complete Secretion of Phaseolin in Transgenic Tobacco Leaves

The results presented above suggest that the vacuolar sorting of phaseolin occurs because of recognition events that are saturable. This is consistent with the presence of a receptor mechanism that recognizes specific structural features of phaseolin. Signals for sorting to the plant vacuoles have been identified on several proteins. There is not a unique consensus; however, a high percentage of hydrophobic residues is a characteristic of some of the known C-terminal vacuolar sorting signals (Bednarek et al., 1990; Neuhaus et al., 1991; Saalbach et al., 1996). When analyzing the phaseolin sequence, we observed that there was an enrichment in hydrophobic amino acids at its C terminus: Lys is followed by the sequence Gly–Ala–Phe–Val–Tyr (Figure 1). Therefore, we tested whether the hydrophobic C-terminal end of phaseolin contains information necessary for vacuolar targeting. We introduced a stop codon after residue 417 of T343F to generate T343FΔ418 (hereafter referred to as Δ418; Figure 1). Thus, Δ418 lacks the last four amino acid residues, Ala–Phe–Val–Tyr, which form a highly hydrophobic sequence. We produced transgenic tobacco plants expressing Δ418 under the control of the cauliflower mosaic virus 35S promoter.

We isolated total proteins from transgenic leaves and analyzed them by SDS-PAGE, which was followed by immuno-

blot analysis with the anti-phaseolin antiserum. The results are shown in Figure 7A. Plants expressing Δ418 accumulated a major phaseolin polypeptide with an M_r of ~45,000 and a minor one with a slightly lower relative molecular mass. These polypeptides are of the relative mass expected for intact phaseolin and may represent the glycosylated and unglycosylated forms of the protein. Immunoblot analysis of an extract from tobacco leaves expressing T343F is also shown in Figure 7A for comparison. Clearly, the typical phaseolin fragmentation products present when the protein is synthesized in transgenic plants are absent in Δ418-expressing plants.

Because the fragments in the 20- to 25-kD range are formed when phaseolin is targeted to the tobacco vacuoles, we investigated whether Δ418 accumulates in the apoplast instead of being delivered to vacuoles. We were able to visualize Δ418 and T343F on silver-stained SDS-PAGE gels, after immunoprecipitation with the anti-phaseolin antiserum (Figure 7B). When phaseolin was immunoprecipitated from total leaf homogenates, Δ418 and the fragmentation products of T343F were clearly detectable (Figure 7B, lanes 1
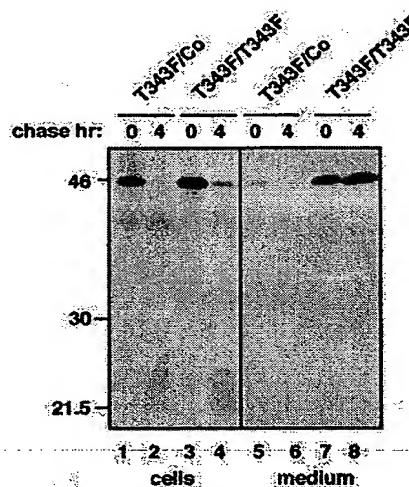
Figure 6. Vacuolar Sorting of T343F in Transgenic Protoplasts Is Saturable.

Protoplasts from leaves of transgenic tobacco expressing T343F were transfected with either the plasmid without an insert (T343F/Co) or the plasmid encoding T343F (T343F/T343F). Cells were subjected to a 1-hr pulse with ^{35}S-methionine and ^{35}S-cysteine and chased for the indicated periods of time. Cell homogenates and the corresponding incubation media were immunoprecipitated with the anti-phaseolin antiserum and then analyzed by SDS-PAGE and fluorography. Numbers at left indicate molecular mass markers in kilodaltons.
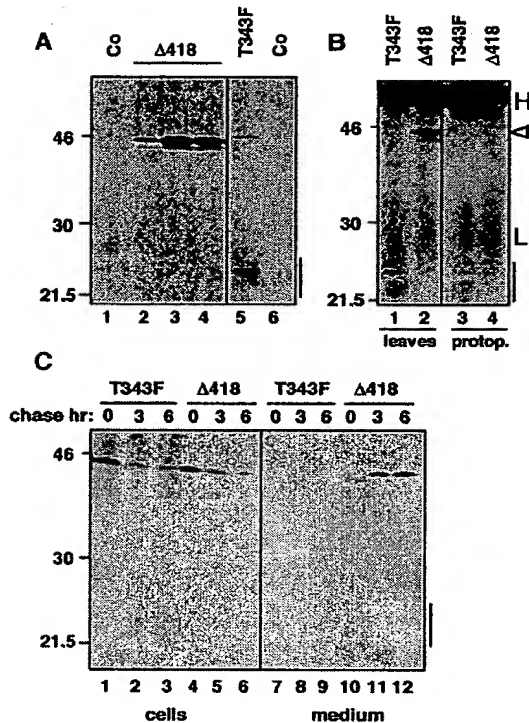
Figure 7. Δ418 Is Completely Secreted from Leaf Cells of Transgenic Tobacco.

(A) Leaves from transgenic tobacco expressing T343F or Δ418 (three independent transgenic plants) or transformed with the plasmid without inserts (Co) were homogenized, and the homogenates were analyzed by SDS-PAGE, followed by protein gel blotting and immunodetection with the anti-phaseolin antiserum. Lanes 1 to 4 and 5 and 6 represent independent immunodetections.

(B) Whole-leaf homogenates or protoplast (protop.) homogenates from plants expressing T343F or Δ418 were immunoprecipitated with the anti-phaseolin antiserum and analyzed by SDS-PAGE. The gels were then stained with silver stain. The arrowhead indicates the position of intact phaseolin. The positions of the heavy (H) and light (L) chains of the antibodies used for immunoprecipitation are indicated.

(C) Protoplasts from transgenic tobacco plants expressing T343F or Δ418 were pulse-labeled for 1 hr with ^{35}S-methionine and ^{35}S-cysteine and chased for the indicated periods of time. Cell homogenates and the corresponding incubation media were immunoprecipitated with the anti-phaseolin antiserum and then analyzed by SDS-PAGE and fluorography.

In (A) to (C), the vertical bar indicates phaseolin fragmentation products. Numbers at left indicate molecular mass markers in kilodaltons.

and 2; there was insufficient intact T343F to be detected, indicating that its proportion with respect to the T343F fragments was overestimated by immunoblot analysis). However, when phaseolin was immunoprecipitated from leaf protoplast preparations, Δ418 was absent, whereas the vacuolar fragments of T343F were, as expected, still present (Figure 7B, lanes 3 and 4). This indicates that Δ418 accumulates in the apoplast.

To study the trafficking of Δ418, we subjected protoplasts isolated from Δ418 transgenic leaves, or from T343F transgenic leaves as a control, to pulse–chase analysis, followed by immunoprecipitation with the anti-phaseolin antiserum, SDS-PAGE, and fluorography (Figure 7C). During the chase, Δ418 was secreted from the cells and accumulated in the extracellular medium; no fragmentation products were detectable. A very small amount of fragmentation products became visible in the intracellular Δ418 immunoprecipitates upon extremely long film exposures, indicating that a very small proportion of the mutated phaseolin can reach the vacuoles and that the protein can be fragmented. This proportion of mutated phaseolin, however, constitutes much less than 1% of total immunoprecipitable phaseolin (data not shown).

The C Terminus of Phaseolin Is Exposed on the Surface of the Protein

The fact that Δ418 is secreted could be due to different causes. (1) The last four amino acids are an essential part of the vacuolar sorting signal for phaseolin, and their deletion directly abolishes recognition by the sorting machinery. And (2), removal of the four residues causes defects to the overall structure of phaseolin, and secretion could then be due to a wrong conformation of the phaseolin molecules; these could be no longer recognized by the vacuolar sorting machinery, because of an altered surface exposure of a sorting signal constituted by amino acids other than the deleted ones. The three-dimensional structure of phaseolin has been established in large part (Lawrence et al., 1990, 1994); however, no information is available for the last 22 amino acids at the C terminus or for a few other segments of the protein.

We devised experiments to assess the oligomerization state of Δ418 phaseolin and the exposure of the C terminus of T343F phaseolin, thereby attempting to distinguish between the two possibilities mentioned above. We transiently transformed protoplasts with plasmids encoding T343F, Δ418, or Δ360 and subjected them to 1 hr of pulse-labeling. We then loaded the cell homogenates onto a 5 to 25% sucrose sedimentation velocity gradient. Phaseolin was then immunoprecipitated from each gradient fraction (Figure 8). The position of Δ418 along the gradient matched the position of the trimeric, transport-competent T343F exactly. In Figure 8, the assembly-defective mutant Δ360 (Pedrazzini et al., 1997) marked the position of phaseolin monomers along
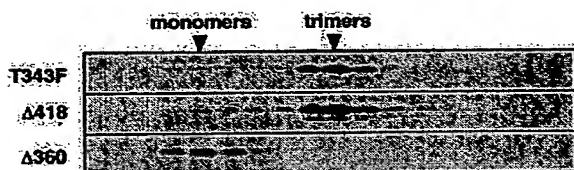
Figure 8. Δ418 Is Trimeric.

Tobacco protoplasts were transiently transfected with plasmids encoding T343F, Δ418, or Δ360 and pulse-labeled for 1 hr with ^{35}S-methionine and ^{35}S-cysteine. Cell homogenates were subjected to sedimentation velocity fractionation on a continuous 5 to 25% [w/v] sucrose gradient. Gradient fractions were subjected to immunoprecipitation with the anti-phaseolin antiserum and analyzed by SDS-PAGE and fluorography. The top of the gradients is at left. Only the portions of the gels containing phaseolin are shown.

the gradient. We thus conclude that Δ418 is competent for rapid and efficient trimerization and indistinguishable by velocity centrifugation from normal T343F phaseolin. Therefore, the absence of the four C-terminal residues does not alter the oligomeric state of phaseolin.

Secretion of T343F during transient expression is accompanied by a late and slight decrease in molecular weight. This can barely be seen when comparing lanes 9 and 10 of Figure 2A or lanes 5 and 6 of Figure 5, but it becomes evident when SDS-PAGE is performed for longer periods of time, allowing more pronounced separation of the different polypeptides, as shown in Figure 9, lanes 1 and 2. When synthesis was allowed to occur in the presence of the inhibitor of N-glycosylation tunicamycin, the time-dependent decrease in the molecular weight of secreted T343F was still evident, indicating that it is not due to glycan trimming (Figure 9, lanes 3 and 4). However, secreted Δ418 did not undergo any decrease in molecular weight during the chase (Figure 9, cf. lanes 5 and 6, and 7 and 8). This strongly suggests that secreted T343F undergoes removal of a few amino acids at the C terminus and that therefore the C-terminal tetrapeptide required for vacuolar sorting is exposed on the surface of the assembled protein. By using SDS-PAGE analysis, we could not establish whether intracellular T343F deposited into the vacuoles also undergoes this processing, because of the major fragmentation event that accompanies vacuolar deposition.

The results presented in Figures 8 and 9 suggest the existence of a direct interaction between the C terminus of phaseolin and the vacuolar sorting mechanism.

DISCUSSION

We have shown here that vacuolar sorting of phaseolin in tobacco leaf cells is very efficient in transgenic plants but can

be saturated when high levels of phaseolin are expressed in transiently transformed protoplasts. Saturation leads to Golgi-mediated secretion of the excess of protein. We have also shown that deletion of the last four C-terminal amino acid residues abolishes sorting of phaseolin to the vacuole and leads to its complete secretion. Therefore, vacuolar sorting of phaseolin, a storage protein of the 7S (vicilin) class, is most likely mediated by a saturable receptor system that recognizes a signal located at the C terminus of its ligand protein.

Vacuolar Sorting of Phaseolin

In the transgenic tobacco plants that we produced, T343F phaseolin was not secreted to detectable levels but almost saturated the vacuolar sorting mechanism: no additional vacuolar phaseolin was detectable upon transient overexpression, and excess phaseolin was secreted. It is indeed possible that these transgenic plants have already slowly adapted to handle the excess of vacuolar protein that they are synthesizing. A further increase in the synthesis of phaseolin could require further adaptation that is not achieved during the short transient expression experiments. This would be a possible explanation for the otherwise fortunate coincidence between the level of phaseolin synthesis and the capacity of its sorting mechanism in transgenic tobacco leaves.

A saturable targeting mechanism is consistent with it being receptor mediated. In yeast, the receptor-mediated vacuolar targeting of proteinase A and carboxypeptidase Y is saturable: when expressed at artificially high levels, these two proteins are secreted (Rothman et al., 1986; Stevens et al., 1986). A putative receptor that recognizes some but not all plant vacuolar targeting signals in vitro has been identified in pea, but it is not known whether this receptor recognizes phaseolin (Kirsch et al., 1996; Paris et al., 1997).
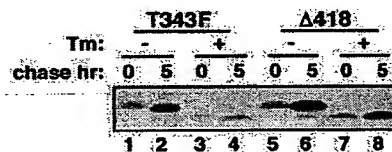
Figure 9. Secreted T343F, but Not Δ418, Undergoes Post-Translational Proteolytic Processing.

Tobacco protoplasts were transiently transfected with plasmids encoding T343F or Δ418 and subjected to pulse–chase in the presence (+) of tunicamycin (Tm) or in the absence (−) of the inhibitor. Pulse was for 1 hr with ^{35}S-methionine and ^{35}S-cysteine, and chase was for 0 or 5 hr. The incubation media were then immunoprecipitated with the anti-phaseolin antiserum and analyzed by SDS-PAGE and fluorography. Note the increase in mobility of T343F, but not of Δ418, between the pulse and the chase.

Transient expression in tobacco protoplasts has been used to study vacuolar sorting of other plant proteins. Targeting of the wild-type forms of barley lectin (Bednarek et al., 1990) and Brazil nut 2S albumin (Saalbach et al., 1996) was not saturated, but recombinant vacuolar chitinase of tobacco showed partial secretion that was directly proportional to its level of expression (Neuhaus et al., 1994). Because more than one targeting mechanism exists for plant vacuoles (Matsuoka et al., 1995), it is possible that not all of them are saturated at similar levels of protein synthesis; the different plasmids used for transfection may also lead to different expression levels, but saturability is clearly not unique to phaseolin.

The saturation of the vacuolar targeting of phaseolin suggests either that a component of the sorting machinery is not sufficiently abundant to handle the number of phaseolin molecules synthesized or that above a certain concentration, phaseolin undergoes a change in conformation that inhibits recognition by the sorting machinery. Electron microscopy shows that in tobacco leaf vacuoles, as in bean storage vacuoles, phaseolin forms protein bodies, probably because of the acidic environment. Currently, however, we have no evidence that upon transient expression, similar structures or aggregates of phaseolin are formed during its transport before being deposited into the vacuole: we were unable to detect any phaseolin aggregates by velocity gradient centrifugation, and the proteolytic trimming of secreted T343F phaseolin strongly suggests that the sequence necessary for its vacuolar sorting remains available for interactions with the sorting machinery.

Further support of this hypothesis comes from digestion of secreted T343F phaseolin with endoglycosidase H, which allowed us to determine that in plant cells, secretion of an overexpressed vacuolar protein occurs via the Golgi-mediated pathway and not via an alternative route. The fact that the vast majority of secreted T343F phaseolin molecules are resistant to endoglycosidase H digestion also indicates that their glycan is accessible to Golgi-processing enzymes: this confirms that at the time of passage through at least some of the Golgi stacks, most of the phaseolin that will be secreted is not aggregated. Therefore, if secretion is due to aggregation, this is a late and transient event during the trafficking of phaseolin, and the aggregates are easily disrupted during analysis. Therefore, we suggest that secretion is instead due to insufficient abundance of the sorting machinery.

The fact that the glycan of T343F acquires a complex structure both in secreted phaseolin and in phaseolin deposited into the vacuole has implications for the definition of the routes that lead to the two different locations as well. Analogous to the targeting to the lytic compartments of other eukaryotes, it is assumed that sorting of vacuolar proteins from secretion occurs in the trans-Golgi network, but this has not been proven directly. Because in plant cells N-linked glycans acquire complex structures in the medial and trans-Golgi complex (Fitchette-Lainé et al., 1994), our data indicate that the two routes do not fork before at least the cis-

and medial Golgi complex stacks have been visited. This confirms the previous observation that complex glycans are present both on secreted and on vacuolar plant proteins (Faye et al., 1989; Vitale and Bollini, 1995) and extends it to the synthesis of an individual protein in cells from a single tissue.

The Sorting Signal

The sequence Ala–Phe–Val–Tyr is uncharged and mainly hydrophobic. Phaseolin is the product of a small gene family, and the tetrapeptide is conserved in both α and β polypeptides (Slightom et al., 1985). Short propeptides at the C terminus are known to determine the vacuolar sorting of tobacco chitinase (seven amino acids; Neuhaus et al., 1991) and Brazil nut 2S albumin (four amino acids; Saalbach et al., 1996). A longer C-terminal propeptide (15 amino acids) contains the vacuolar sorting signal of barley lectin (Bednarek et al., 1990). A vacuolar sorting consensus sequence cannot be derived, but these propeptides are also rich in hydrophobic amino acids. Mutational analysis revealed the importance of hydrophobic residues in barley lectin (Dombrowski et al., 1993) but not in chitinase (Neuhaus et al., 1994). It should also be considered that in the native plants, the 7S and 2S storage proteins and barley lectin accumulate in storage vacuoles, whereas chitinase is located in vegetative vacuoles. Transport from the Golgi complex to the two types of vacuoles involves different structures (Chrispeels, 1983; Hohl et al., 1996). Eventually, in mature mesophyll cells, the two types of vacuoles probably merge (Schroeder et al., 1993; Paris et al., 1996), but it is possible that the mechanisms recognizing the C-terminal vacuolar sorting signals are more than one and are distinct for vegetative and storage proteins.

When T343F is secreted during transient expression, the C terminus is most probably trimmed by hydrolases that also are secreted by the protoplasts, indicating that it is exposed on the surface of phaseolin trimers. Whether this trimming also occurs upon delivery to the vacuole cannot be established by SDS-PAGE analysis, because vacuolar phaseolin is fragmented in tobacco cells. However, phaseolin accumulated in the storage vacuoles of bean cotyledonary cells lacks a few amino acids with respect to the ER-associated precursor (Bollini et al., 1982; D'Amico et al., 1992). The N-terminal sequence of phaseolin purified from bean cotyledons is the one predicted after signal peptide removal, indicating that trimming must be at the C terminus (Paaren et al., 1987). Therefore, the Ala–Phe–Val–Tyr sequence probably constitutes or is part of a short transient peptide and is most likely exposed on the surface of phaseolin trimers. This makes the sequence a good candidate for direct interactions with the vacuolar sorting machinery and does not favor the alternative hypothesis that deletion of the sequence affects the overall conformation of phaseolin and inhibits recognition by the sorting machinery only as an indirect effect.

Secretion is a safe haven for phaseolin: indeed, Δ418 accumulates in its intact form. When wild-type phaseolin is expressed in transgenic plants, the default targeting to the vacuole leads to its proteolytic cleavage in all tissues tested, with the only exception being rice endosperm (Murai et al., 1983; Sengupta-Gopalan et al., 1985; Bagga et al., 1992; Zheng et al., 1995; Pedrazzini et al., 1997). Comparisons of the amounts of T343F and Δ418 accumulated in leaves are beyond the scope of this work, but our data suggest that secretion could be a very effective means to preserve phaseolin from degradation. This could be considered as a more general, useful strategy when planning the expression of heterologous proteins of biotechnological interest in the plant secretory pathway. A strategy that has been successful in enhancing the stability of foreign vacuolar proteins is their accumulation in the ER via the KDEL/HDEL system normally used by soluble ER residents (Wandelt et al., 1992; Tabe et al., 1995; Khan et al., 1996). It would be interesting to compare accumulation in the ER versus the apoplast by using the same passenger proteins.

Vacuolar Sorting and ER Quality Control

The fact that Δ418 phaseolin is not sorted to the vacuole but is secreted means that the assembly-defective phaseolin constructs Δ363 and Δ360 (Pedrazzini et al., 1997), which are the results of more extensive C-terminal deletions, do not possess the vacuolar targeting signal of trimeric phaseolin as well. On the contrary, HiMet phaseolin, a trimeric but unstable mutated phaseolin that is transported to the vacuole and rapidly degraded there, still bears the sequence Ala–Phe–Val–Tyr (Hoffman et al., 1988; Pueyo et al., 1995). Assembly-defective phaseolin is subjected to ER quality control: it is extensively retained in the ER, where it shows prolonged interactions with the chaperone BiP and then is slowly degraded (Pedrazzini et al., 1997). The location where assembly-defective phaseolin is degraded is not known. Possible candidates are a subcompartment of the ER or the cytosol after dislocation from the ER; another candidate is the vacuole, which would be reached by a Golgi-independent pathway (Pedrazzini et al., 1997). The results we have presented here indicate that in whichever compartment degradation occurs, targeting of assembly-defective phaseolin to degradation by quality control does not depend on the signal that normally sorts phaseolin to the hydrolytic compartments of the plant cell.

METHODS

Recombinant DNA

The strategies to construct T343F and Δ360 phaseolin have been described previously (Pedrazzini et al., 1997). To construct Δ418, the antisense oligonucleotide 5'-CTGCAG<u>TCA</u>ACCCTTTCTTCCCTT-

TTGC-3' was used in polymerase chain reaction to introduce a stop codon (underlined) after residue 417 in the T343F phaseolin coding sequence. The construct thus lacks the last four residues of wild-type phaseolin (Ala, Phe, Val, and Tyr). Δ418 was introduced into the vector pDHA (Tabe et al., 1995) for transient expression experiments. For constitutive expression in transgenic tobacco, pDHA containing Δ418 was inserted into the HindIII site of the binary vector pGA470 (An et al., 1985), which was then used to transform *Agrobacterium tumefaciens* LBA4404 by electroporation.

Production of Transgenic Tobacco Plants and Transient Transformation of Leaf Protoplasts

The transformed agrobacterium was used to produce transgenic tobacco (*Nicotiana tabacum* cv Petit Avana SR1) plants, as described by Pedrazzini et al. (1997).

Protoplasts were prepared from axenic leaves (4 to 7 cm long) of untransformed tobacco or from transgenic tobacco expressing T343F or Δ418 phaseolin. Protoplasts were subjected to polyethylene glycol–mediated transfection, as described previously (Pedrazzini et al., 1997). Vector pDHA without inserts was used as a negative control for transfection. Unless otherwise stated, 40 μg of plasmid was used to transform 10^6 protoplasts in 1 mL. After transfection, protoplasts were allowed to recover overnight in the dark at 25°C in K3 medium (Pedrazzini et al., 1997), at a concentration of 10^6 cells/mL, before performing pulse–chase experiments.

Pulse–Chase Labeling of Protoplasts and Analysis of Phaseolin

Pulse–chase labeling of protoplasts using Pro-Mix (a mixture of ^{35}S-methionine and ^{35}S-cysteine; Amersham) was performed as described previously (Pedrazzini et al., 1997). In some experiments, before radioactive labeling, protoplasts were incubated for 45 min at 25°C in K3 medium supplemented with 50 μg/mL tunicamycin (stock solution 5 mg/mL in 10 mM NaOH, stored at 4°C; Boehringer Mannheim) and maintained in the presence of the inhibitor for the entire experiment. At the desired pulse–chase time points, 3 volumes of W5 medium (154 mM NaCl, 5 mM KCl, 125 mM $CaCl_2 \cdot 2H_2O$, and 5 mM glucose) was added, and the samples were centrifuged for 5 min at 50g. The supernatant, containing the proteins secreted into the incubation medium, was removed, leaving 100 μL to cover the protoplasts. The supernatant and protoplasts were frozen separately in liquid nitrogen and stored at −80°C. Homogenization of the frozen samples and immunoprecipitation of phaseolin using rabbit antiserum raised against phaseolin purified from bean cotyledons were performed as described previously (Pedrazzini et al., 1997). Unless otherwise stated, for each pulse–chase time point, immunoprecipitation was performed using cell homogenate and incubation medium from the same amount of protoplasts. Digestion of immunoprecipitated proteins with endoglycosidase H (recombinant; Boehringer Mannheim) was performed as described previously (Ceriotti et al., 1991).

For the analysis of phaseolin assembly by sedimentation velocity, after radioactive labeling and homogenation, homogenates were brought to 8 mM $MgCl_2$ and 3 mM ATP and loaded on top of a continuous 5 to 25% [w/v] linear sucrose gradient made in 150 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, and 50 mM Tris-Cl, pH 7.5. Samples were centrifuged at 39,000 rpm in an SW40 Ti rotor (Beckman Instruments, Inc., Fullerton, CA) for 20 hr at 20°C. Phaseolin was then immunoselected from each gradient fraction.

SDS-PAGE analysis was always performed on 15% acrylamide gels. Rainbow ^{14}C-methylated proteins (Amersham) were used as molecular weight markers. Gels were treated with 2,5-diphenyloxazole dissolved in dimethyl sulfoxide and radioactive polypeptides revealed by fluorography. Quantification of the relative intensities of bands in the fluorographs was performed by microdensitometry, using a Camag TLC Scanner II (Camag, Muttenz, Switzerland). Care was taken to use film exposures that were in the linear range of film darkening.

Total Leaf Protein Extraction and Protein Gel Blot Analysis

For analysis of total proteins, leaves from transformed plants were homogenized in an ice-cold mortar with 0.2 M NaCl, 1 mM EDTA, 2% β-mercaptoethanol, 0.2% Triton X-100, and 0.1 M Tris-Cl, pH 7.8, supplemented with Complete (Boehringer Mannheim) protease inhibitor cocktail. The buffer/tissue ratio was 6:1 (milliliters per gram fresh weight of tissue). The homogenate was centrifuged at 1500 g at 4°C for 10 min, and the supernatant was separated by SDS-PAGE. Immunodetection was performed using the enhanced chemiluminescence (ECL) system (Amersham), following the manufacturer's protocol. Proteins were detected using the anti-phaseolin antiserum.

To determine whether Δ418 phaseolin is located intracellularly or extracellularly, we immunoselected protoplasts or leaf extracts from transgenic plants with the anti-phaseolin antiserum, and the selected material was analyzed by SDS-PAGE followed by silver staining.

Immunomicroscopy

Leaf fragments (1 to 2 mm) were fixed in 0.8% glutaraldehyde and 3.3% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4, at 4°C for 2 hr and in 1% OsO$_4$ in the same buffer for 2 hr. They were then dehydrated in an ethanol series and embedded in Spurr's resin (Spurr, 1969). Thin sections were etched with 1% aqueous periodic acid and incubated overnight at 4°C in the presence of either the anti-phaseolin antiserum, diluted 1:1000 with PBS-BSA-C (Aurion, Wageningen, The Netherlands), or preimmune serum as control. After washing, the samples were incubated for 1 hr at room temperature in goat anti-rabbit 15-nm gold complex (1:20; Biocell, Cardiff, UK) and stained with 2% uranyl acetate and lead citrate.

ACKNOWLEDGMENTS

REFERENCES

Ahmed, S.U., Bar-Peled, M., and Raikhel, N.V. (1997). Cloning and subcellular location of an Arabidopsis receptor-like protein that shares common features with protein-sorting receptors of eukaryotic cells. Plant Physiol. 114, 325–336.

An, G., Watson, B.D., Stachel, S., Gordon, M.P., and Nester, E.W. (1985). New cloning vehicles for transformation of higher plants. EMBO J. 4, 277–284.

Bagga, S., Sutton, D., Kemp, J.D., and Sengupta-Gopalan, C. (1992). Constitutive expression of the β-phaseolin gene in different tissues of transgenic alfalfa does not ensure phaseolin accumulation in non-seed tissue. Plant Mol. Biol. 19, 951–958.

Bednarek, S.Y., Wilkins, T.A., Dombrowski, J.A., and Raikhel, N.V. (1990). A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. Plant Cell 2, 1145–1155.

Bollini, R., Van der Wilden, W., and Chrispeels, M.J. (1982). A precursor of the reserve protein, phaseolin, is transiently associated with the endoplasmic reticulum of developing Phaseolus vulgaris cotyledons. Physiol. Plant. 55, 82–92.

Ceriotti, A., Pedrazzini, E., Fabbrini, M.S., Zoppè, M., Bollini, R., and Vitale, A. (1991). Expression of wild-type and mutated vacuolar storage protein phaseolin in Xenopus oocytes reveals relationships between assembly and intracellular transport. Eur. J. Biochem. 202, 959–968.

Chrispeels, M.J. (1983). The Golgi apparatus mediates the transport of phytohemagglutinin to the protein bodies in bean cotyledons. Planta 158, 140–151.

D'Amico, L., Valsasina, B., Daminati, M.G., Fabbrini, M.S., Nitti, G., Bollini, R., Ceriotti, A., and Vitale, A. (1992). Bean homologs of the mammalian glucose regulated proteins: Induction by tunicamycin and interaction with newly synthesized storage proteins in the endoplasmic reticulum. Plant J. 2, 443–455.

Denecke, J., Gosselé, V., Botterman, J., and Cornelissen, M. (1989). Quantitative analysis of transiently expressed genes in plant cells. Methods Mol. Cell. Biol. 1, 19–27.

Denecke, J., Botterman, J., and Deblaere, R. (1990). Protein secretion in plant cells can occur via a default pathway. Plant Cell 2, 51–59.

Dombrowski, J.E., Schroeder, M.R., Bednarek, S.Y., and Raikhel, N.V. (1993). Determination of the functional elements within the vacuolar targeting signal of barley lectin. Plant Cell 5, 587–596.

Dyer, J.M., Nelson, J.W., and Murai, N. (1995). Extensive modifications for methionine enhancement in the β-barrels do not alter the structural stability of the bean seed storage protein phaseolin. J. Prot. Chem. 14, 665–678.

Faye, L., Johnson, K.D., Sturm, A., and Chrispeels, M.J. (1989). Structure, biosynthesis, and function of asparagine-linked glycans on plant glycoproteins. Physiol. Plant. 75, 309–314.

Fitchette-Lainé, A.-C., Gomord, V., Chekkafi, A., and Faye, L. (1994). Distribution of xylosylation and fucosylation in the plant Golgi apparatus. Plant J. 5, 673–682.

Hoffman, L.M., Donaldson, D.D., and Herman, E.M. (1988). A modified storage protein is synthesized, processed, and degraded in the seeds of transgenic plants. Plant Mol. Biol. 11, 717–729.

Hohl, I., Robinson, D.G., Chrispeels, M.J., and Hinz, G. (1996). Transport of storage proteins to the vacuole is mediated by vesicles without a clathrin coat. J. Cell Sci. 109, 2539–2550.

Hunt, D.C., and Chrispeels, M.J. (1991). The signal peptide of a vacuolar protein is necessary and sufficient for the efficient secretion of a cytosolic protein. Plant Physiol. 96, 18–25.

Khan, M.R.I., Ceriotti, A., Tabe, L., Aryan, A., McNabb, W., Moore, A., Craig, S., Spencer, D., and Higgins, T.J.V. (1996). Accumulation of a sulphur-rich seed albumin from sunflower in the leaves of transgenic subterranean clover (Trifolium subterraneum L.). Transgen. Res. 5, 1–7.

Kirsch, T., Saalbach, G., Raikhel, N.V., and Beevers, L. (1996). Interaction of a potential vacuolar targeting receptor with amino- and carboxyl-terminal targeting determinants. Plant Physiol. 111, 469–474.

Lawrence, M.C., Suzuki, E., Varghese, J.N., Davis, P.C., Van Donkelaar, A., Tulloch, P.A., and Colman, P.M. (1990). The three-dimensional structure of the seed storage protein phaseolin at 3 Å resolution. EMBO J. 9, 9–15.

Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J., and Colman, P.M. (1994). Structure of phaseolin at 2.2 Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. J. Mol. Biol. 238, 748–776.

Matsuoka, K., Bassham, D.C., Raikhel, N.V., and Nakamura, K. (1995). Different sensitivity to wortmannin of two vacuolar sorting signals indicates the presence of distinct sorting machineries in tobacco cells. J. Cell Biol. 130, 1307–1318.

Murai, N., Sutton, D.W., Murray, M.G., Slightom, J.L., Merlo, D.J., Reichert, N.A., Sengupta-Gopalan, C., Stock, C.A., Barker, R.F., Kemp, J.D., and Hall, T.C. (1983). Phaseolin gene from bean is expressed after transfer to sunflower via tumor-inducing plasmid vectors. Science 222, 476–482.

Neuhaus, J.-M., Sticher, L., Meins, F., Jr., and Boller, T. (1991). A short C-terminal sequence is necessary and sufficient for targeting of chitinases to the plant vacuole. Proc. Natl. Acad. Sci. USA 88, 10362–10366.

Neuhaus, J.-M., Pietrzak, M., and Boller, T. (1994). Mutation analysis of the C-terminal vacuolar targeting peptide of tobacco chitinase: Low specificity of the sorting system, and gradual transition between intracellular retention and secretion into the extracellular space. Plant J. 5, 45–54.

Okita, T.W., and Rogers, J.C. (1996). Compartmentation of proteins in the endomembrane system of plant cells. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 327–350.

Paaren, H.E., Slightom, J.L., Hall, T.C., Inglis, A.S., and Blagrove, R.J. (1987). Purification of a seed glycoprotein: N-terminal and deglycosylation analysis of phaseolin. Phytochemistry 26, 335–343.

Paris, N., Stanley, C.M., Jones, R.L., and Rogers, J.C. (1996). Plant cells contain two functionally distinct vacuolar compartments. Cell 85, 563–572.

Paris, N., Rogers, S.W., Jiang, L., Kirsch, T., Beevers, L., Phillips, T.E., and Rogers, J.C. (1997). Molecular cloning and further characterization of a probable plant vacuolar sorting receptor. Plant Physiol. 115, 29–39.

Pedrazzini, E., Giovinazzo, G., Bielli, A., de Virgilio, M., Frigerio, L., Pesca, M., Faoro, F., Bollini, R., Ceriotti, A., and Vitale, A. (1997). Protein quality control along the route to the plant vacuole. Plant Cell 9, 1869–1880.

Pueyo, J.J., Chrispeels, M.J., and Herman, E.M. (1995). Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole. Planta 196, 586–596.

Rothman, J.H., Hunter, C.P., Valls, L.A., and Stevens, T.H. (1986). Overproduction-induced mislocalization of a yeast vacuolar protein allows isolation of its structural gene. Proc. Natl. Acad. Sci. USA 83, 3248–3252.

Saalbach, G., Rosso, M., and Schumann, U. (1996). The vacuolar targeting signal of the 2S albumin from Brazil nut resides at the C terminus and involves the C-terminal propeptide as an essential element. Plant Physiol. 112, 975–985.

Schroeder, M.R., Borkhsenious, O.N., Matsuoka, K., Nakamura, K., and Raikhel, N.V. (1993). Colocalization of barley lectin and sporamin in vacuoles of transgenic tobacco plants. Plant Physiol. 101, 451–458.

Sengupta-Gopalan, C., Reichert, N.A., Barker, R.F., Hall, T.C., and Kemp, J.D. (1985). Developmentally regulated expression of the bean β-phaseolin gene in tobacco seed. Proc. Natl. Acad. Sci. USA 82, 3320–3324.

Slightom, J.L., Drong, R.F., Klassy, R.C., and Hoffman, L.M. (1985). Nucleotide sequences from phaseolin cDNA clones: The major storage proteins from Phaseolus vulgaris are encoded by two unique gene families. Nucleic Acids Res. 13, 6483–6498.

Spurr, A.R. (1969). A low-viscosity epoxy resin embedding medium for electron microscopy. J. Ultrastruct. Res. 26, 31–43.

Stevens, T.H., Rothman, J.H., Payne, G.S., and Scheckman, R. (1986). Gene dosage-dependent secretion of yeast vacuolar carboxypeptidase Y. J. Cell Biol. 102, 1551–1557.

Sturm, A., Van Kuik, J.A., Vliegenthart, J.F.G., and Chrispeels, M.J. (1987). Structure, position, and biosynthesis of the high mannose and the complex oligosaccharide side chains of the bean storage protein phaseolin. J. Biol. Chem. 262, 13392–13403.

Tabe, L.M., Wardley-Richardson, T., Ceriotti, A., Aryan, A., McNabb, W., Moore, A., and Higgins, T.J.V. (1995). A biotechnological approach to improving the nutritive value of alfalfa. J. Anim. Sci. 73, 2752–2759.

Vitale, A., and Bollini, R. (1995). Legume storage proteins. In Seed Development and Germination, J. Kigel and G. Galili, eds (New York: Marcel Dekker), pp. 73–102.

Wandelt, C.I., Khan, M.R.I., Craig, S., Schroeder, H.E., Spencer, D., and Higgins, T.J.V. (1992). Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. Plant J. 2, 181–192.

Zheng, Z., Sumi, K., Tanaka, K., and Murai, N. (1995). The bean seed storage protein β-phaseolin is synthesized, processed, and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. Plant Physiol. 109, 777–786.

δ-Tonoplast intrinsic protein defines unique plant vacuole functions

GUANG-YUH JAUH*, ANDREAS M. FISCHER†, HOWARD D. GRIMES†, CLARENCE A. RYAN, JR.*, AND JOHN C. ROGERS*‡

*Institute of Biological Chemistry, Box 6340, and †Department of Genetics and Cell Biology/Botany, Washington State University, Pullman, WA 99164

ABSTRACT Plant cell vacuoles may have either storage or degradative functions. Vegetative storage proteins (VSPs) are synthesized in response to wounding and to developmental switches that affect carbon and nitrogen sinks. Here we show that VSPs are stored in a unique type of vacuole that is derived from degradative central vacuoles coincident with insertion of a new tonoplast intrinsic protein (TIP), δ-TIP, into their membranes. This finding demonstrates a tight coupling between the presence of δ-TIP and acquisition of a specialized storage function and indicates that TIP isoforms may determine vacuole identity.

Uniquely, in contrast to yeast or mammalian cells, plant cells may contain separate vacuoles with protein storage and digestive functions (1, 2). It is not known how functionally distinct vacuoles are generated or maintained. Plant vacuole tonoplast membranes contain abundant tonoplast intrinsic proteins (TIPs) that may function as aquaporins (3), but the quantities present seem to be far in excess of what is required for water transport (4). Protein storage vacuoles (PSVs), containing seed-type storage proteins, are marked by the presence of α-TIP, and lytic or degradative vacuoles (LVs) are marked by the presence of γ-TIP (1, 2). These observations indicate that a specific TIP isoform correlates with a specific vacuole function (5). A further test of this hypothesis would require demonstration that a third functionally distinct vacuole carried a different TIP isoform. Here we define a third functionally distinct vacuole in plant cells and demonstrate that it is marked specifically by δ-TIP.

MATERIALS AND METHODS

Antibodies and Immunocytochemistry. Anti-α-TIP protein antiserum (6) and VM23 antiserum to purified γ-TIP protein from radish root (7) were generously provided by M. Chrispeels (6) and M. Maeshima (7), respectively. Synthetic peptides were synthesized and antisera to the peptides coupled to keyhole limpet hemocyanin were generated by Quality-Controlled Biochemicals, Hopkinton, MA. For antibody purification, the peptides were coupled via an amino-terminal Cys residue to SulfoLink agarose (Pierce) according to the manufacturer's instructions, and peptide-specific antibodies were affinity-purified as described previously (8) for use in all procedures. Antisera to proteinase inhibitor I (Inh I) and II (Inh II) have been described previously (9–11). Fluorochrome-tagged secondary antibodies were purchased from Jackson ImmunoResearch. Plant tissues were fixed in either formaldehyde/acetic acid/ethyl alcohol or 3.7% paraformaldehyde, and paraffin-embedded sections were prepared as described (12). After removal of paraffin and rehydration, the sections were blocked as described (2). The double-label protocol to identify two different rabbit antibodies separately with immunofluorescence has been described (2). Briefly, the first primary antibody was completely covered by incubation with an excess of anti-rabbit F(ab')₂ fragment coupled to lissamine rhodamine before incubating with the second primary antibody and detection with a fluorescein isothiocyanate (FITC)-coupled secondary antibody. Slides were viewed and photographed under epifluorescence with an Olympus BX-50 microscope containing a multifilter set (no. 61002, Chroma Technology, Brattleboro VT) that permits simultaneous viewing of blue (462 nm), green (531 nm), and red (627 nm) emissions that are excited at 402 nm, 496 nm, and 571 nm, respectively. This filter set converts most background fluorescence from plant tissues, including that from chlorophyll, to gray, and thus facilitates identification of fluorescence resulting from specific antibodies. Photographic prints were scanned into a computer, and images representing individual red (lissamine rhodamine) and green (FITC) emissions were generated by using ADOBE PHOTOSHOP software (Adobe Systems, Mountain View, CA).

Plant Material. Growth of soybean plants and depodding to induce vegetative storage protein (VSP) accumulation has been described (13–15). Tomato, petunia, and tobacco plants were grown in Washington State University greenhouses, and other plant materials were purchased from a local supermarket. Tissue extracts for Western blot analyses (16) were prepared by homogenization at 25°C in 0.0625 M Tris·HCl, pH 6.8/2% SDS with a Polytron-type homogenizer (PRO250; PRO Scientific, Monroe, CT). After homogenization the extracts were allowed to sit for 1 hr at 25°C and then clarified by centrifugation at 12,000 × g.

RESULTS

Previous studies to identify specific TIP isoforms in vacuoles used antiserum to α-TIP, a single protein purified from bean seeds (6) and VM23 antiserum recognizing a purified γ-TIP protein purified from radish root (7). We noted that the carboxyl-terminal, cytoplasmic-tail sequences of the different TIPs varied such that it was possible to make antibodies specific for each: γ-TIP, THEQLPTTDY (barley, GenBank accession no. X80266; rice, D25534; *Arabidopsis*, Z34662 and X72581; radish, D84669); α-TIP, HQPLATEDY (*Phaseolus*, X62873) and HQPLAPEDY (*Arabidopsis*, X63551); DIP (17), HTPLPTSEDYA (*Antirrhinum*, X70417) and HAPLPTSEDYA (potato, U65700; tobacco, X54855); and δ-TIP, HV-PLASADF (*Arabidopsis*, U39485). Accordingly, rabbits were immunized with the following peptides coupled to keyhole limpet hemocyanin: γ-TIP, CSRTHEQLPTTDY; α-TIP, C(aminohexanoic acid)HQPLAPEDY; DIP, CHTPLPT-

Abbreviations: VSP, vegetative storage protein; TIP, tonoplast intrinsic protein; PSV, protein storage vacuole; LV, lytic vacuole; Inh I, protease inhibitor I; Inh II, protease inhibitor II; PVM, paraveinal mesophyll; ΔV, delta vacuole; FITC, fluorescein isothiocyanate; DIC, differential interference contrast.
‡To whom reprint requests should be addressed. e-mail: bcjroger@wsu.edu.

SEDYA; and δ-TIP, CHVPLASADF, and affinity-purified antibodies were tested on dot-blots containing the different peptides coupled to BSA. Each antibody preparation was specific, having at least a 100-fold-higher affinity for its corresponding peptide than for any of the other peptides (data not presented). Use of the anti-DIP peptide antibodies will be described elsewhere.

Anti-TIP Peptide Antibody Specificity. The specificity of the antibodies is demonstrated by the following results from Western blot analyses of different plant tissue extracts (Fig. 1). The anti-α-TIP protein antiserum (Fig. 1*A*) identified a band of the appropriate 26-kDa size in only the pea root tip extract (lane 1), and a similar pattern was obtained with our anti-α-TIP peptide antibodies (Fig. 1*B*). The VM23 anti-γ-TIP protein antibodies identified appropriate bands in both pea root tip and radish root (lane 2) (Fig. 1*D*), a pattern that was indistinguishable from that obtained with our anti-γ-TIP peptide antibodies (Fig. 1*E*). Therefore, results with the anti-peptide antibodies were indistinguishable from results in which antibodies that detect the entire α- and γ-TIP proteins were used. In immunofluorescence studies with pea and barley root tip cells (2), the anti-α- and anti-γ-TIP antibodies label separate α-TIP- and γ-TIP-containing vacuoles, respectively (G.-Y.J. and J.C.R., unpublished data); therefore, the ≈26-kDa band detected by the anti-γ-TIP antibodies in pea root tip extracts (Fig. 1*D* and *E*, lane 1) is γ-TIP and does not represent cross-reactivity with α-TIP. [The ≈40-kDa bands represent TIP dimers that form when tissue extracts are heated during preparation of samples for electrophoresis (6, 7). Additionally, the variation in size of the ≈26-kDa TIP monomers is a result of proteolytic cleavage of the first transmembrane domain from the rest of the protein in certain plant tissues (18).] In contrast, the δ-TIP anti-peptide antibodies (Fig. 1*C*) identified specific TIP bands in pea root tips, radish root, flower petals from petunia (lane 3) and tobacco (lane 4), and potato tubers (lane 5). Neither α-TIP nor γ-TIP were detected in the extracts from flower petals and potato tuber. These findings led us to ask which functions might be associated with δ-TIP in potato tuber or flower petal vacuoles.

Many plant species accumulate VSPs (19, 20). VSPs have physical characteristics that differ substantially from seed-type storage proteins present in PSVs and include proteins with diverse functions such as protease inhibitors, acid phosphatases, and lipoxygenases. VSPs may accumulate during vegetative growth, serving as temporary sites of carbon and nitrogen storage, and are mobilized as an energy source during seed development (19). In other circumstances, VSPs such as

protease inhibitors accumulate in response to pathogen attack or wounding (9). Protease inhibitors are abundant in potato tubers, but initially are stored in leaves in plantlets until rhizomes form; then the inhibitors disappear from leaves and begin to accumulate in the developing tubers (21).

Immunofluorescence Localization. We used immunofluorescence on paraffin-embedded tissue sections to assess localization of VSPs and their association with δ-TIP in several different tissues of potato, tomato, petunia, and soybean plants by using specific antibodies. In potato tuber (Fig. 2), strong labeling with the anti-δ-TIP antibodies (red) was present on central vacuole tonoplast in most cells (open arrow), and the same vacuoles were strongly labeled with antibodies to potato protease inhibitor I (21) (Inh I, green). Occasional cells had two distinct vacuoles: in addition to large vacuoles containing δ-TIP and Inh I (asterisk) there were smaller vacuoles (solid arrow) that did not label with either antibody. Additionally, in these cells, both δ-TIP and Inh I appeared to be present in numerous small organelles within the cytoplasm.

Proteinase inhibitors in tomato leaves are induced by wounding or by methyl jasmonate treatment (10). We used antibodies to tomato proteinase inhibitor II (Inh II) (11) to compare its localization to that of δ-TIP in sections of leaves from plants treated with methyljasmonate for 2 days to induce accumulation of Inh II (Fig. 3 *a* and *b*). On Western blots, tomato leaf extracts have no detectable α- or γ-TIP, but have abundant δ-TIP (data not presented). Both Inh II (Fig. 3, green) and δ-TIP (red) antibodies strongly labeled epidermal cells (Fig. 3*a*), bundle sheath cells (not shown), and cells (Fig. 3*b*) with a tissue distribution typical of paraveinal mesophyll (PVM) as characterized in legumes. The two antibodies colocalized in small vacuoles (solid arrows) that were separate from the central vacuole in the cells; neither antigen appeared to be associated with the central vacuole. Controls used in all experiments support the specificity of patterns shown: (*i*) single labeling with each of the antibodies gave the same cell and organelle staining pattern (data not presented); and (*ii*) when sections labeled with the δ-TIP antibodies were incubated with the excess of rhodamine-conjugated anti-rabbit IgG F(ab′)₂ secondary antibodies, washed, and then incubated with the FITC-conjugated anti-rabbit IgG secondary antibodies, no green labeling was obtained (data not presented). This demonstrates that the F(ab′)₂ antibodies completely blocked all of the δ-TIP antibody sites so that FITC labeling in the sections shown in Fig. 3 must be a result of the presence of anti-proteinase inhibitor II antibodies at those sites.
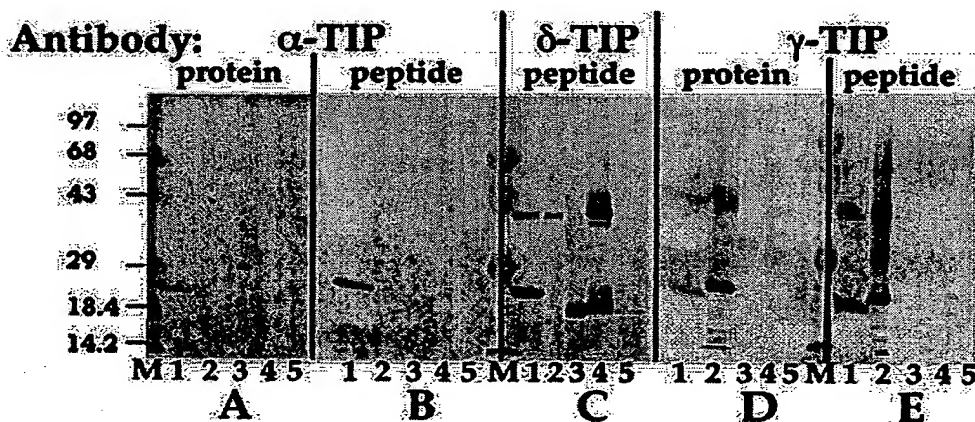

FIG. 1.   Western blot analyses of plant tissue extracts with different anti-TIP antibodies. Membranes carrying 300 μg of protein from extracts of pea root tips (1), radish root (2), petunia flower petals (3), tobacco flower petals (4), and potato tuber (5) were tested with antibodies to α-TIP protein (*A*) (8) (1:1,000), α-TIP C-terminal peptide (*B*) (0.2 μg/ml), δ-TIP peptide (*C*) (0.95 μg/ml), γ-TIP protein (*D*) [VM23 antiserum (7), 1:1,000], and γ-TIP peptide (*E*) (0.24 μg/ml). M, molecular mass markers; sizes (in kDa) are indicated to the left.

Cell Biology: Jauh *et al.*

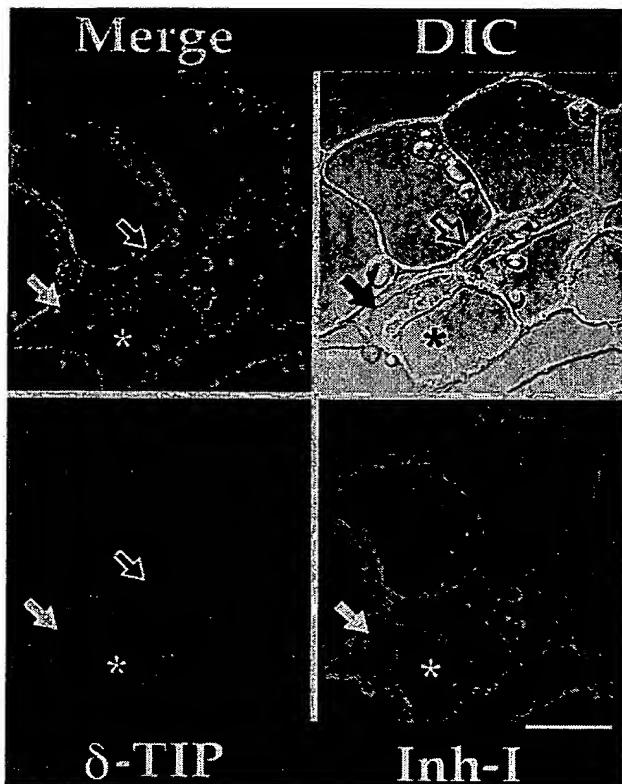*Proc. Natl. Acad. Sci. USA 95 (1998)*   12997

FIG. 2. Immunofluorescence on potato tuber. A differential interference contrast (DIC) image of the potato tuber section is shown at the upper right. The double-label image (Merge) is in the upper left, with separated anti-δ-TIP (used at 19 μg/ml) (red) and anti-protease inhibitor I (Inh I) (1:100) (green) images at the bottom. Open arrow, central vacuole tonoplast staining with anti-δ-TIP, where Inh I is closely adherent; solid arrow, vacuole with tonoplast not staining with anti-δ-TIP and not containing Inh I in a cell with a larger vacuole (asterisk) labeled with the two antibodies. (Bar = 50 μm.)

We found that tomato flower petals contain high concentrations of protease inhibitors I and II, 94 and 524 μg/ml of tissue extract, respectively. When sections of tomato flower petals were probed with antibodies against δ-TIP and Inh II, anti-δ-TIP (Fig. 3c, red) strongly labeled the tonoplast of the central vacuole (solid arrow) in all epidermal cells; double-labeling demonstrated that Inh II (Fig. 3c, green) was present in the contents of these vacuoles (open arrow). Anti-δ-TIP antibodies also labeled the central vacuole tonoplast of epidermal cells in petunia flower petals (Fig. 4a). As shown in Fig. 4b, pigment is present in the central vacuole of the upper epidermal cells.

To test further the apparent association of δ-TIP and vacuoles containing VSPs, we utilized a soybean system in which deposition of VSPs in vacuoles of epidermal, bundle sheath, and PVM cells is induced by continuous depodding to remove the major nitrogen sink (13, 19, 22). Control plants were grown under identical conditions but were allowed to develop pods (13). Western blots demonstrated γ- and δ-TIP in leaf extracts of both control and depodded plants (data not presented). Double-label immunofluorescence of leaf sections with the two antibodies demonstrated that γ-TIP (Fig. 5a, red) labeled central vacuole tonoplast of epidermal (arrow), bundle sheath, and PVM (white dot) cells with no difference between control (Fig. 5a Left) and depodded (Fig. 5a Right) plants. In contrast, the pattern of δ-TIP labeling changed substantially in leaves from control versus depodded plants. In control (green, Fig. 5a Left), little labeling of epidermal (arrow) or PVM
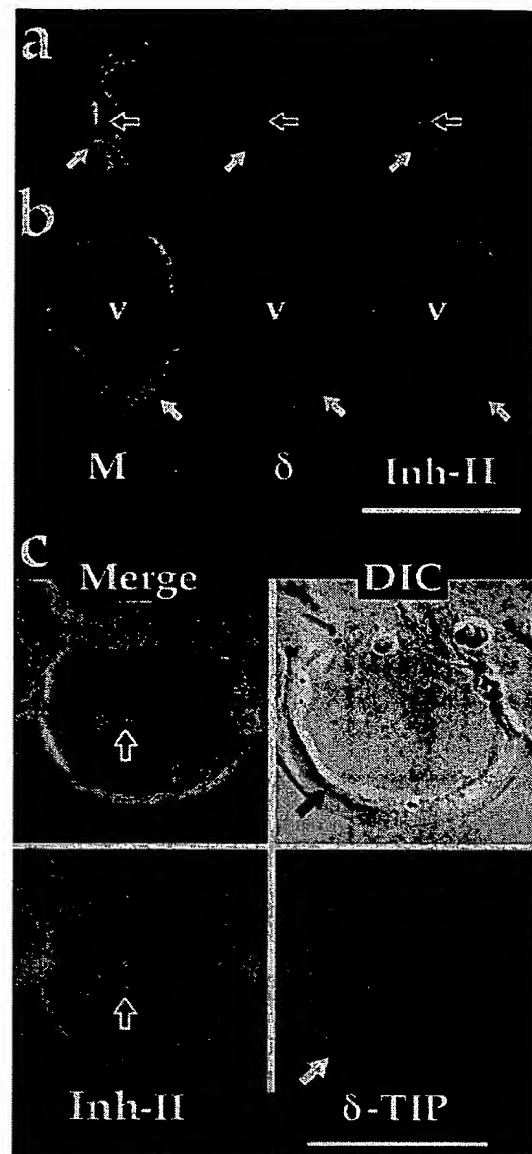
FIG. 3. Immunofluorescence on tomato tissues. Anti-δ-TIP (δ) is in red and anti-Inh II (used at 1:100) is in green. M, double-label image. (a) Epidermal cell from leaf section. Solid arrow indicates small vacuole stained with both antibodies; open arrow indicates central vacuole. (b) PVM cell. Solid arrow as for a; v, central vacuole. (c) Epidermal cell from bottom surface of flower petal. Solid arrow indicates central vacuole tonoplast stained with anti-δ-TIP; open arrow indicates aggregated Inh II protein within lumen of central vacuole. (Bars = 25 μm.)

(white dot) was observed, while prominent labeling of those cells was observed in depodded plants (Fig. 5a Right). This is a specific difference because the maximum intensity of labeling of cells in the palisade layer (P) was similar in both control and depodded samples. The reproducibility of these observations was tested in a blinded manner by providing unlabeled slides with anti-γ- and anti-δ-TIP double-labeled sections to three independent observers; each observer accurately identified control and depodded samples. In Fig. 5b, the difference in δ-TIP abundance is compared in sections labeled with only that antibody. Spongy mesophyll cells stained with similar intensity in leaves from both controls and depodded plants. Clearly,
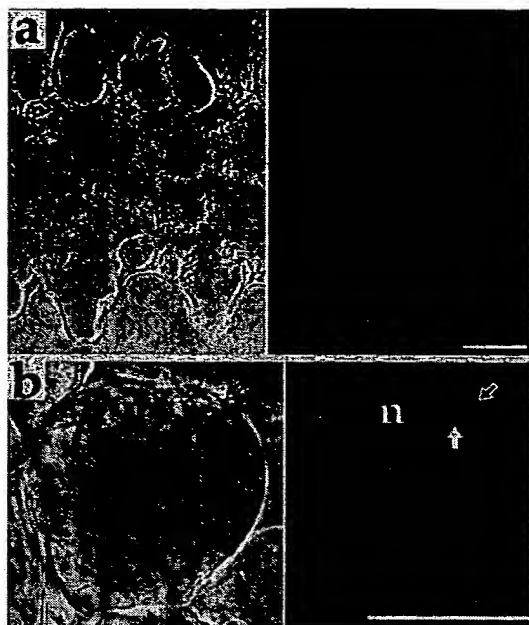
FIG. 4.   Immunofluorescence on petunia petal. (*a*) DIC (*Left*) and anti-δ-TIP labeling (*Right*). Upper and lower epidermal cells are shown at top and bottom, respectively. (*b*) Living cell from upper epidermis of petal. DIC image (*Left*) and endogenous fluorescence from vacuole contents (*Right*). Solid arrow, limit of vacuole; open arrow, cell wall; n = nucleus. (Bars = 25 μm.)

however, there was abundant staining of the tonoplast in PVM (dot) and epidermal (asterisk) in leaves from depodded plants while very little staining of these cells was evident in the control plants. When similar sections were double-labeled with anti-δ-TIP antibodies and antibodies to soybean VSPs, the VSPs were present in PVM and epidermal cells (13, 15) whose vacuoles now carried δ-TIP in their tonoplast (data not presented). Thus, induction of VSP storage in soybean leaf cell vacuoles was accompanied by induction of δ-TIP expression in tonoplast of the same vacuoles while expression of γ-TIP was unchanged.

DISCUSSION

These data demonstrate that vacuoles whose tonoplast is labeled with anti-δ-TIP antibodies are used by plant cells to store pigments and VSPs, proteins synthesized in response to developmental and environmental cues. Central vacuoles in petunia and tomato flower petal epidermal cells contain pigments and protease inhibitors. Tonoplast in most central vacuoles in potato tubers labeled strongly with anti-δ-TIP antibodies, and only those vacuoles contained Inh I. Vacuoles in tomato leaf epidermal, bundle sheath, and PVM cells labeled with anti-δ-TIP antibodies and contained Inh II; these vacuoles were smaller and separate from central vacuoles in the same cells. In soybean leaves, central vacuoles marked by γ-TIP in their tonoplast acquired abundant δ-TIP in the same tonoplast when storage of VSPs in the vacuoles was induced by depodding. This finding demonstrates a tight coupling between the presence of δ-TIP and acquisition of a specialized vacuolar storage function by vacuoles that otherwise would have been assumed to have a lytic function given the presence of γ-TIP in their membranes. These results in aggregate define "delta vacuoles" (ΔVs) with δ-TIP in their tonoplast as specialized storage organelles that are structurally and functionally distinct from PSVs and LVs.
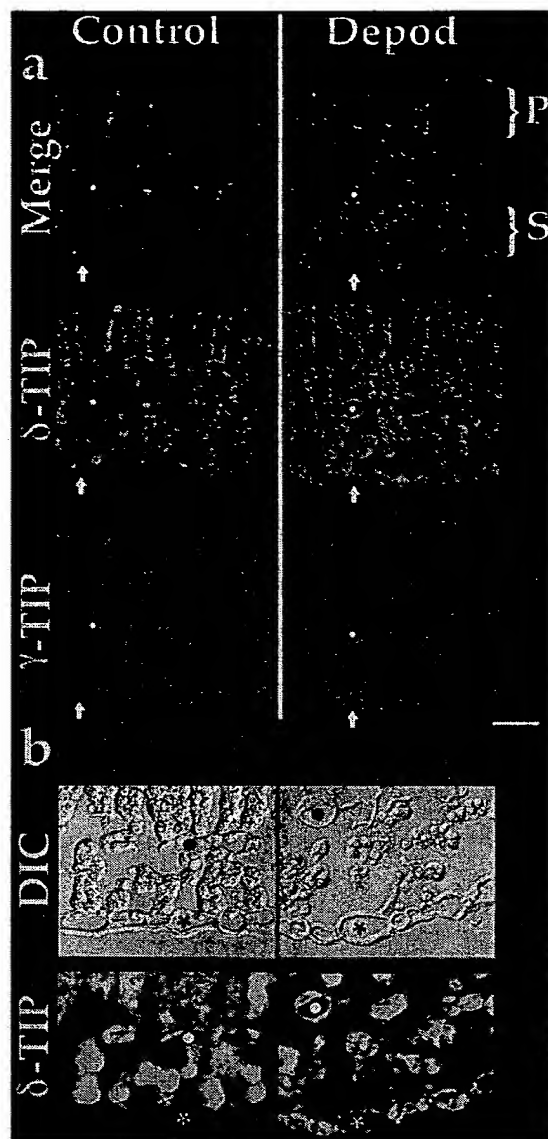
FIG. 5.   Immunofluorescence on soybean leaf. (*a*) Sections from leaves harvested at 4 weeks from control (*Left*) and continuously depodded (*Right*) plants are shown. Indicated are the palisade cell (P) and spongy mesophyll cell (S) layers; the white dot indicates a PVM cell, while the arrow indicates an epidermal cell on the lower surface of the leaf. Anti-δ-TIP is in green, and anti-γ-TIP (used at 8.8 μg/ml) is in red, with the double-label (Merge) image at the top. (Bar = 50 μm.) (*b*) Sections from leaves of control and depodded plants as in *a*. (*Upper*) DIC image. (*Lower*) Single label with anti-δ-TIP. ✱, Epidermal cell; ●, PVM cell.

The pattern of labeling observed with anti-δ-TIP antibodies was unusual. Frequently large and numerous aggregates were present along the tonoplast and within the vacuole lumen (e.g., Fig. 2). In leaves from depodded soybean plants, these δ-TIP-containing aggregates frequently appeared almost to fill the vacuole lumen (e.g., Fig. 5*b*, cells adjacent to asterisk). This appearance indicates that formation of ΔVs may involve more than insertion of δ-TIP molecules into preexisting tonoplast. Perhaps these aggregates represent newly synthesized membrane with high concentrations of δ-TIP that accumulates as vesicles or strands within the vacuole lumen. It is likely that the aggregates provide a clue as to how the internal vacuole environment is changed to permit VSP accumulation.

Cell Biology: Jauh *et al.*

Proc. Natl. Acad. Sci. USA 95 (1998)   12999

Mechanisms for sorting soluble proteins to ΔVs have not been elucidated. The finding, however, that soybean vacuoles carrying the LV marker γ-TIP may be converted to ΔVs indicates that traffic of soluble proteins to at least some ΔVs probably involves a vacuolar-sorting receptor associated with the LV pathway (23, 24). This would explain why the vacuolar-sorting determinant of prosporamin (25), a sweet potato tuber protease inhibitor (26), interacts with that receptor (23, 27).

1. Hoh, B., Hinz, G., Jeong, B.-K. & Robinson, D. G. (1995) *J. Cell Sci.* **108**, 299–310.
2. Paris, N., Stanley, C. M., Jones, R. L. & Rogers, J. C. (1996) *Cell* **85**, 563–572.
3. Chrispeels, M. J. & Maurel, C. (1994) *Plant Physiol.* **105**, 9–13.
4. Higuchi, T., Hisada, H., Morishima, S., Okada, Y. & Maeshima, M. (1998) *Plant Cell Physiol.*, in press.
5. Neuhaus, J.-M. & Rogers, J. C. (1998) *Plant Mol. Biol.,* **38**, 127–144.
6. Johnson, K. D., Herman, E. M. & Chrispeels, M. J. (1989) *Plant Physiol.* **91**, 1006–1013.
7. Maeshima, M. (1992) *Plant Physiol.* **98**, 1248–1254.
8. Holwerda, B. C., Galvin, N. J., Baranski, T. J. & Rogers, J. C. (1990) *Plant Cell* **2**, 1091–1106.
9. Ryan, C. A. & Huisman, O. C. (1967) *Nature (London)* **214**, 1047–1049.
10. Ryan, C. A. (1980) *Curr. Top. Cell. Regul.* **17**, 1–23.
11. Graham, J. S., Pearce, G., Merrywheather, J., Titani, K., Ericson, L. & Ryan, C. A. (1985) *J. Biol. Chem.* **260**, 6561–6564.
12. Berlyn, G. P. & Miksche, J. P. (1976) *Botanical Microtechnique and Cytochemistry* (Iowa State Univ. Press, Ames).
13. Tranbarger, T. J., Franceschi, V. R., Hildebrand, D. F. & Grimes, H. D. (1991) *Plant Cell* **3**, 973–987.
14. Bunker, T. W., Loetje, D. S., Stephenson, L. C., Creelman, R. A., Mullet, J. E. & Grimes, H. D. (1995) *Plant Cell* **7**, 1319–1331.
15. Stephenson, L. C., Bunker, T. W., Dubbs, W. E. & Grimes, H. D. (1998) *Plant Physiol.* **116**, 923–933.
16. Rogers, S. W., Burks, M. & Rogers, J. C. (1997) *Plant J.* **11**, 1359–1368.
17. Culianez-Macia, F. A. & Martin, C. (1993) *Plant J.* **4**, 717–725.
18. Inoue, K., Takeuchi, Y., Nishimura, M. & Hara-Nishimura, I. (1995) *Plant Mol. Biol.* **28**, 1089–1101.
19. Staswick, P. E. (1994) *Annu. Rev. Plant Physiol. Plant Mol Biol.* **45**, 303–322.
20. Herman, E. M. (1994) in *Advances in Structural Biology,* ed. Malhotra, S. (JAI Press, Greenwich, CT), pp. 243–283.
21. Ryan, C. A. & Pearce, G. (1998) *Annu. Rev. Cell Dev. Biol.* **14**, in press.
22. Staswick, P. E. (1988) *Plant Physiol.* **87**, 250–254.
23. Kirsch, T., Paris, N., Butler, J. M., Beevers, L. & Rogers, J. C. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3403–3407.
24. Paris, N., Rogers, S. W., Jiang, L., Kirsch, T., Beevers, L., Phillips, T. E. & Rogers, J. C. (1997) *Plant Physiol.* **115**, 29–39.
25. Nakamura, K., Matsuoka, K., Mukumoto, F. & Watanabe, N. (1993) *J. Exp. Bot.* **44**, Suppl., 331–338.
26. Yeh, K.-W., Chen, J.-C., Lin, M.-I., Chen, Y.-M. & Lin, C.-Y. (1997) *Plant Mol. Biol.* **33**, 565–570.
27. Kirsch, T., Saalbach, G., Raikhel, N. V. & Beevers, L. (1996) *Plant Physiol.* **111**, 469–474.

Isolation and characterization of a cDNA clone encoding wheat germ agglutinin

(lectin/wheat)

NATASHA V. RAIKHEL AND THEA A. WILKINS

Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, MI 48824-1312

ABSTRACT   Two sets of synthetic oligonucleotides coding for amino acids in the amino- and carboxyl-terminal portions of wheat germ agglutinin were synthesized and used as hybridization probes to screen cDNA libraries derived from developing embryos of tetraploid wheat. The nucleotide sequence for a cDNA clone recovered from the cDNA library was determined by dideoxynucleotide chain-termination sequencing in vector M13. The amino acid sequence deduced from the DNA sequence indicated that this cDNA clone (pNVR1) encodes isolectin 3 of wheat germ agglutinin. Comparison of the deduced amino acid sequence of clone pNVR1 with published sequences indicates isolectin 3 differs from isolectins 1 and 2 by 10 and 8 amino acid changes, respectively. In addition, the protein encoded by pNVR1 extends 15 amino acids beyond the carboxyl terminus of the published amino acid sequence for isolectins 1 and 2 and includes a potential site for N-linked glycosylation. Utilizing the insert of pNVR1 as a hybridization probe, we have demonstrated that the expression of genes for wheat germ agglutinin is modulated by exogenous abscisic acid. Striking homology is observed between wheat germ agglutinin and chitinase, both of which are proteins that bind chitin.

Lectins, sugar-binding proteins derived mainly from plant sources, have been of great value as specific probes for investigating the distribution and function of carbohydrates on the surfaces of animal cells (1, 2). In recent years, however, the notion has become widely accepted that the ability of lectins to distinguish discrete sugars did not arise fortuitously during evolution (2), and as a result, there has been increased interest in the synthesis and biochemistry of this group of proteins. Wheat germ agglutinin (WGA), the first cereal lectin characterized in detail, binds specifically to the sugar N-acetylglucosamine and to chitin, a polymer of N-acetylglucosamine residues (3, 4). In the hexaploid wheat *Triticum aestivum*, WGA exists as three closely related isolectins derived from the A, B, and D genomes (5, 6). Comparison of the amino acid sequences for isolectin 1 (A genome) and isolectin 2 (D genome) indicates that these proteins differ at four residues (7, 8). The amino acid sequence for isolectin 3 (B genome), the least abundant form, is not yet available. These three isolectins randomly associate into functional dimers *in vivo* (5) and are immunologically indistinguishable (9).

In wheat plants, WGA is found in the embryos and adventitious roots (9–11). During embryogenesis, WGA expression is under temporal control (12). Accumulation of WGA is tissue-specific and cell-type specific in various organs of the embryo (e.g., coleoptile, coleorhiza, and radicle) (9, 10). In other species of Gramineae, a lectin immunologically related to WGA is synthesized during seed development and in the roots of adult plants (13, 14).

Furthermore, the accumulation of lectin is modulated by the hormone abscisic acid (12, 15). Although biochemical, immunological, and microscopic studies have helped to characterize the composition and distribution of WGA (3–8, 10, 11), the genes for WGA have not been isolated.

We are interested in investigating the molecular mechanisms that regulate the developmental tissue-specific expression of WGA genes. To isolate clones for WGA, cDNA libraries from developing grains of the tetraploid wheat *Triticum durum* (AABB) were used. Here, we report the isolation and the nucleotide sequence of a cDNA clone that we presume to encode isolectin 3.* Using this clone as a hybridization probe, we present evidence that the expression of WGA genes is modulated by abscisic acid. Because of the common ability of WGA and chitinase to bind chitin, we searched for amino acid homology using the recently published sequence for chitinase from *Phaseolus vulgaris* (16). We found strong homology between the amino terminus of chitinase and four regions of WGA. The significance of this similarity is addressed.

MATERIALS AND METHODS

Plant Material. Wheat *T. aestivum* L. (AABBDD) cv. Marshall was obtained from the Minnesota Crop Improvement Association (St. Paul, MN). Plants were grown as previously described (11), and embryos were collected at 20 days after bloom (anthesis) according to Raikhel and Quatrano (12). Abscisic acid treatment involved culturing isolated embryos in the dark at 27°C for 3 days on filter paper containing growth medium (15) with and without 10^{-4} M abscisic acid (Sigma).

Materials. Two cDNA libraries, derived from mRNA isolated from developing wheat grains of *T. durum* (AABB) cv. Mexicali at 3 and 4 weeks post-anthesis, were provided by C. Brinegar (ARCO Plant Cell Research Institute, Dublin, CA). Two sets of degenerate synthetic oligonucleotides were prepared for amino acid regions in isolectin 1 (8): for the sequence Asn-Met-Glu-Cys-Pro-Asn-Asn in the amino-terminal region (residues 9–15), probe 2, TTR TAC CTY ACR GGN TTR TT; and for the sequence Cys-Thr-Asn-Asn-Tyr-Cys-Cys in the carboxyl terminal region (residues 141–147), probe 1, ACR TGN TTR TTR ATR ACR AC. The oligonucleotide mixtures were synthesized on an Applied Biosystems (Foster City, CA) 380 DNA synthesizer by a solid-phase method (17) and separated by electrophoresis on a 20% polyacrylamide gel containing 8 M urea in TBE, pH 8.3 (0.89 M Tris/0.089 M boric acid/2.7 mM EDTA). The oligonucleotides were eluted in 0.5 M ammonium acetate/10 mM magnesium acetate/0.1% NaDodSO₄, and then end-labeled with ^{32}P using T4 polynucleotide kinase (18).

Abbreviation: WGA, wheat germ agglutinin.
*This sequence of isolectin 3 of wheat germ agglutinin is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J02961).

Isolation and Screening of cDNA Clones. The cDNA libraries, in *Escherichia coli* strain DH5α (19), were plated directly onto nitrocellulose filters laid on agar plates containing Luria broth medium with ampicillin at 50 μg/ml (20). After colonies were established, the bacteria were lysed, and the filters were probed with oligonucleotide probes 1 and 2 as follows. The temperature of hybridization (T_H) for each oligonucleotide was calculated using the formula $T_H = T_D - 12°C$, where $T_D = 2°C ×$ (the number of A·T base pairs) plus $4°C ×$ (the number of G·C base pairs). Replicate filters were prehybridized in 6× SSC (0.9 M sodium chloride and 0.09 M sodium citrate, pH 7.0) (1× SSC = 0.15 M sodium chloride/0.015 M sodium citrate, pH 7) plus 0.25% nonfat milk, and hybridized in the same buffer containing the labeled oligonucleotide probes and 0.1% NaDodSO₄ at 36°C (probe 1) or 38°C (probe 2). After hybridization, filters were washed three times in 6× SSC/0.25% nonfat milk/0.1% NaDodSO₄ at room temperature for 10 min, followed by a 2-min wash at 46°C (probe 1) or at 48°C (probe 2). Filters were dried and autoradiographed for 16–18 hr. Colonies that produced positive signals were selected and rescreened using the same probes under the same conditions.

DNA Sequence Determination. Inserts from recombinant plasmids were purified by electrophoresis in low-melting-point agarose. Excised cDNA inserts or appropriate restriction fragments were then subcloned into M13mp18 or M13mp19. Dideoxynucleotide chain-termination sequencing from single-stranded M13 templates was accomplished using a Bethesda Research Laboratories M13 sequencing kit with the exception that dGTP was replaced by 7-deaza-2'-deoxyguanosine triphosphate (Boehringer Mannheim).

RNA Blot Analysis. Total RNA was isolated as described (21) and poly(A)⁺ RNA was purified by chromatography on oligo(dT)-cellulose (18). Poly(A)⁺ RNA was electrophoresed in adjacent lanes (1 μg per lane) on 2% agarose gels containing 6% formaldehyde and then transferred to nitrocellulose (22). Filters were hybridized with inserts labeled by the random primer method of Feinberg and Vogelstein (23) and washed under stringent conditions as described in Thomas (22).

RESULTS

Isolation of cDNA Clones. Two synthetic oligonucleotides, each consisting of 20 nucleotides complementary to the 5' and 3' ends of the coding portion of isolectin 1 mRNA (8), were used for isolation of cDNA clones specific for WGA. These two sequences corresponded to amino acids 9–15 (probe 2) and 141–147 (probe 1). Because of the degeneracy of the sequences involved, probe 2 was a mixture of 64 sequences, and probe 1 was a mixture of 128 sequences. One cDNA clone, pNVR1 [1.0 kilobase (kb)], was selected by hybridization to both probes on the assumption that this insert contains sequences spanning the coding region delimited by the oligonucleotide probes. A second clone, pNVR2 (0.7 kb), was recognized by probe 1 only and is presumably truncated at the 5' end. The restriction map and partial sequence of pNVR2 indicate that it is a shorter version of pNVR1. When the insert from clone pNVR1 was labeled by the random primer method (23) and used as probe to rescreen the cDNA libraries, no additional cDNA clones were retrieved.

Nucleotide Sequence. The cDNA insert of pNVR1 was subcloned into M13mp18 and M13mp19 according to the strategy shown in Fig. 1, and its nucleotide sequence was determined as described. The nucleotide sequence of the cDNA clone and the deduced amino acid sequence are shown in Fig. 2. Clone pNVR1 contains a 558-nucleotide open reading frame encoding a 186-amino acid polypeptide rich in cysteine and glycine but lacking an ATG start codon at the 5' end. Protein sequence analysis indicates that the amino terminus of WGA is blocked (7, 8) so that the first residue (glutamine) of the published sequence may not be the amino terminus of mature WGA. It is,
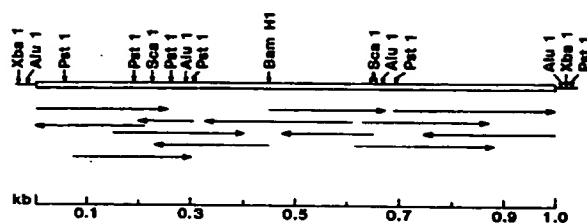

FIG. 1.   Restriction map and sequencing strategy for WGA cDNA clone pNVR1. Open bar, cloned cDNA; arrows, length and direction of the sequenced restriction fragments. Scale of the map is in kb pairs.

therefore, presumably fortuitous that the cDNA clone pNVR1 and the published amino acid sequence for WGA initiate with the same amino acid. The hydropathy plot (24) of the polypeptide encoded by clone pNVR1 shows the polypeptide to be comprised mostly of hydrophilic amino acids (Fig. 3). The polypeptide encoded by pNVR1 extends 15 amino acids beyond the carboxyl terminus of the amino acid sequence published for isolectins 1 and 2 (Fig. 2, squares) (7, 8). The carboxyl-terminal segment contains the most hydrophobic portion of the entire protein (Fig. 3). A potential site for N-linked glycosylation occurs at residues 180–182 (Asn-Ser-Thr) (Fig. 2, dots above squares). The 3'-untranslated region contains four in-frame termination codons (TGA, TGA, TAA, and TAG, underlined in Fig. 2) and a potential polyadenylylation signal (AATAAT, double-underlined in Fig. 2), and terminates in a poly(A) tail that begins 229 nucleotides downstream from that signal.

Comparison with Published Sequences of Isolectins 1 and 2. The amino acid sequence deduced from the cDNA nucleotide sequence (Fig. 2) was compared with published protein sequence data. Re-evaluation of the discrepancies at positions 134 and 150 (Fig. 2, arrows) has indicated a low yield of lysine in addition to glycine for residue 134 (C. Wright, personal communication) and has confirmed the presence of tryptophan at residue 150 (25). The deduced amino acid sequence of pNVR1 was found to differ from the published sequence of isolectin 1 (8) at 10 positions and isolectin 2 at 8 positions (7) (Table 1).

RNA Blot Analysis. Embryos isolated from hexaploid wheat at 20 days post-anthesis were cultured in the presence and absence of abscisic acid (Fig. 4). Equal amounts of poly(A)⁺ RNA from the embryos were fractionated by agarose-formaldehyde gel electrophoresis and transferred to nitrocellulose filters. A 1.1-kb mRNA was detected (Fig. 4) after hybridization with pNVR1 insert labeled by the random primer method (23). The autoradiograph showed that the level of RNA in abscisic acid-treated embryos was several times higher than the level in embryos cultured in the absence of abscisic acid.

Nucleotide and Amino Acid Homology Between WGA and Chitinase. The deduced amino acid sequence of cDNA clone pNVR1 was used to search for homology with chitinase, an enzyme that catalyzes the hydrolysis of 1,4-β linkages of N-acetylglucosamine polymers in chitin. The amino acid homology matrix between clone pNVR1 and chitinase from *P. vulgaris* is shown in Fig. 5. This matrix was generated using the analysis program of Pustell and Kafatos (26) with parameters set so that each letter within the matrix represents a match of 50% or greater over a span of 21 amino acids. Extensive homology between the amino terminus of chitinase and four regions of WGA is apparent.

DISCUSSION

In this paper we present the amino acid sequence of WGA as deduced from a cDNA clone designated pNVR1. That this clone encodes WGA has been verified by comparison of the

```
                                      Asn                                            20
          gln arg cys gly glu gln gly ser gly met glu cys pro asn asn leu cys cys ser gln tyr gly tyr cys gly met

          CAA AGG TGC GGC CAG CAG GGC AGC GGC ATG CAG TGC CCC AAC AAC CTG TGC TGC AGC CAG TAC GGC TAC TGC GGG ATG


                                   Asp            40
          gly gly asp tyr cys gly lys gly cys gln asn gly ala cys trp thr ser lys arg cys gly ser gln ala gly gly

          GGC GGC GAT TAC TGC GGC AAG GGC TGC CAG AAC GGC GCG TGC TGG ACC ACC AAG CGG TGT GGC AGC CAG GCC GGC GGC


          Ala        *        * 60                            *
          lys thr cys pro asn asn his cys cys ser gln tyr gly his cys gly phe gly ala glu tyr cys gly ala gly cys

          AAG ACG TGC CCC AAC AAC CAC TGC TGC AGC CAG TAC GGC CAC TGC GGC TTC GGC GCG GAG TAC TGC GGC GCC GGC TGC


             80                                     Ser                100
          gln gly gly pro cys arg ala asp ile lys cys gly ser gln ala gly gly lys leu cys pro asn asn leu cys cys

          CAG GGC GGC CCC TGC CGC CCC GAC ATC AAG TGC GGC AGC CAG GCC GGC GGC AAG CTG TGC CCC AAC AAC CTC TGC TGC


                  Ser                               Gly 120      Ser
          ser gln trp gly tyr cys gly leu gly ser glu phe cys gly glu gly cys gln asn gly ala cys ser thr asp lys

          AGC CAG TGG GGC TAC TGC GGC CTC GGT TCC GAG TTC TGC GGC GAG GGC TGC CAG AAC GGC GCT TGC AGC ACC GAC AAG


                  ↓                  140                                            ↓
          pro cys gly lys asp ala gly gly arg val cys thr asn asn tyr cys cys ser lys trp gly ser cys gly ile gly

          CCG TGT GGC AAG GAC GCC GGC GGC AGG GTT TGC ACT AAC AAC TAC TGC TGT AGC AAG TGG GGA TCC TGT GGC ATC GGT


              160                                      Ala  ■  ■ ■  ■  ■ ■  ■ ■  ■  . . .
          pro gly tyr cys gly ala gly cys gln ser gly gly cys asp gly val phe ala glu ala ile ala thr asn ser thr

          CCC GGC TAC TGC GGT GCA GGC TGC CAG AGC GGC GGC TGC GAT GGT GTC TTC GCC GAG GCC ATC GCC ACC AAC TCC ACT


          ■  ■ ■ ■
          leu leu ala glu end

          CTT CTC GCA GAA TGA TGA TCTTGCTAATGGTAGTATTGCAACGACGAATAATCCGTGGCAGTTCCATTGCCACGTACGCTTTTCCCTTCACTTAC

          TTTTAGTACTAGTACTTAATAATTCTCTAGCTTGCAATATGCACATGCAGGTTACTGCAGCCACAAACAAAATATTGCTGTAGTACATGCATGGAAATATTACGG

          ACAGAAGTGTTGTGTGGCAATATAGAGTGTACTATAGTTGCCACAAATTAGTGTTCTTGTTATGACCTGCTGTCAGGATGCATGCATGCCTGTTGTAAT

          GTTGGAGTACTTCGTGATATCGTTGCAATATATTACCATGAGTCTCACATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 2. Nucleotide sequence of WGA cDNA clone pNVR1. The deduced amino acid residues are shown above the nucleotide triplets. The differences between the deduced amino acid sequence and the published amino acid sequence of isolectin 2 are designated by the amino acid codes above the deduced sequence. The additional differences with isolectin 1 are designated by asterisks. Proline at position 56 is substituted with threonine, and the histidines at positions 59 and 66 are substituted with glutamine and tyrosine, respectively. Previously described differences at positions 150 (25) and 134 (arrows) have been resolved (C. Wright, personal communication). Four termination codons (single underline) and a putative polyadenylylation signal (double underline) are indicated. Fifteen amino acids that extend beyond the carboxyl terminus of the published sequence for WGA are designated by squares. A potential glycosylation site is indicated by dots above the squares.

deduced amino acid sequence with the sequence determined by direct amino acid sequencing of the purified protein (7, 8). The length of the polypeptide derived from the deduced amino acid sequence is 186 amino acids, and the calculated M_r is 18,754. Nevertheless, pNVR1 does not represent the complete coding sequence for WGA. First, the initiating methionine codon is absent from the cDNA. Second, because WGA is synthesized on and translocated across the rough endoplasmic membrane (27), an amino-terminal signal peptide would be expected (28). Third, there may be one or more amino acids at the amino terminus that have not been detected by peptide sequencing because of blockage of the amino terminus (7, 8). The size of the mRNA recognized by clone pNVR1 predicts a full-length cDNA of 1.1 kb.

The DNA sequence of pNVR1 encodes a protein that extends 15 amino acids beyond the carboxyl terminus of the published amino acid sequence and includes a potential site for N-linked glycosylation. Mature WGA is not a glycoprotein, but its precursor form is glycosylated (27). The site of glycosylation probably lies in the 15 amino acid carboxyl-terminal sequence because the only possible site for glycosylation resides in this region. The glycosylated precursor is known to be processed (27) and to accumulate in protein bodies or vacuoles (10, 29). The WGA precursor in the endoplasmic reticulum-associated fraction is 5 kDa larger than the mature WGA (27). The difference in molecular mass between the precursor and mature WGA may be a consequence of the extra 15 amino acids and glycosylation of the
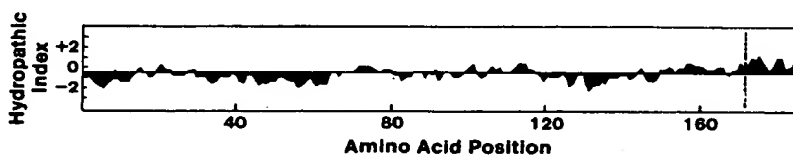

FIG. 3. Hydropathy plot of the protein encoded by cDNA pNVR1. Ordinate, hydropathic index (24); abscissa, amino acid position. The additional 15 amino acids at the carboxyl terminus are right of the broken line.

Table 1.   Amino acids at positions in which there are differences between the residues of isolectins 1 and 2 and the protein encoded by pNVR1

Amino acid	Isolectin 1	Isolectin 2	pNVR1
56	Thr	Pro	Pro
59	Gln	His	His
66	Tyr	His	His
93	Ala	Ser	Ala
9	Asn	Asn	Gly
37	Asp	Asp	Asn
53	Ala	Ala	Lys
109	Ser	Ser	Tyr
119	Gly	Gly	Glu
123	Ser	Ser	Asn
171	Ala	Ala	Gly

carboxyl terminus. The hydropathy plot of the amino acid sequence derived from pNVR1 clearly indicates that the carboxyl terminus of the cloned WGA sequence consists of hydrophobic amino acids, which is consistent with the possibility that it is removed post-translationally. Removal of carboxyl-terminal residues was seen during maturation of napin, a rapeseed storage protein (30). It was recently shown that the lectin concanavalin A (Con A), which is not a glycoprotein, is synthesized as a glycosylated precursor (31). Normal transport of this protein is dependent on the presence of the glycan (32). It is interesting that WGA precursor is a biologically active lectin (27), whereas precursor for Con A does not have lectin activity (31). In other words, the loss of the pro-WGA carboxyl-terminal domain does not relate to its ability to bind N-acetylglucosamine. Alternatively, cleavage of the carboxyl terminus may occur during the purification of WGA such that the mature protein actually contains 186 amino acids *in vivo*.

Clone pNVR1 mRNA contains four termination codons and a 3'-untranslated region. A potential polyadenylylation signal (AAUAAU) is found in the noncoding region followed by a poly(A) tail. Whereas the consensus sequence for the polyadenylylation signal is very highly conserved in animal systems (AAUAAA), plant mRNAs frequently deviate from this theme (33). The deduced amino acid sequence confirms extensive interdomain homology. The 7-amino acid sequence Gly-Cys-Gln-Asn-Gly-Ala-Cys is found at residues 34–40 and again at residues 120–126. Short repeated stretches of
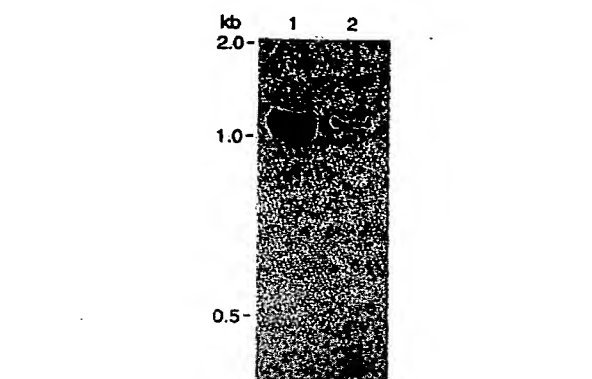
FIG. 4.   RNA blot analysis of WGA mRNA levels. Poly(A)⁺ RNA (1 μg), isolated from embryos excised at 20-day post-anthesis and cultured in the presence (lane 1) and absence (lane 2) of abscisic acid, was separated on a 2% agarose, 6% formaldehyde gel. After transferring the RNA to nitrocellulose, the filter was hybridized to a ³²P-labeled DNA insert from pNVR1 under stringent conditions. Positions of DNA M_r markers were obtained from the ethidium bromide-stained portion of the gel.

Tyr-Cys-Gly, Ala-Gly-Gly, Gly-Cys-Gln, Cys-Cys-Ser, or Cys-Gly-Gly are found throughout the polypeptide.

Amino acid sequence studies on wheat isolectins 1 (A genome) and 2 (D genome) (5–8) indicate that they differ distinctly in their histidine content: two histidines in isolectin 2 and no histidine in isolectin 1 (8). Because clone pNVR1 was isolated from a cDNA library derived from the tetraploid wheat *T. durum* (AABB), the cDNA clone cannot encode isolectin 2 derived from the D genome. Furthermore, pNVR1 probably does not encode isolectin 1 from the A genome. Isolectin 1 does not contain any histidine, whereas pNVR1 encodes a protein containing two histidine residues. Thus, pNVR1 probably represents isolectin 3 derived from the B genome. Although the amino acid compositions of isolectins 2 and 3 are very similar, eight discrete differences were identified between them. At least four of these differences (residues 9, 53, 93, and 119) are authentic. The x-ray crystallographic data for these four positions in isolectin 2 are definitive, and there is no evidence for heterogeneity in peptide preparations (7). The discrepancies at the remaining
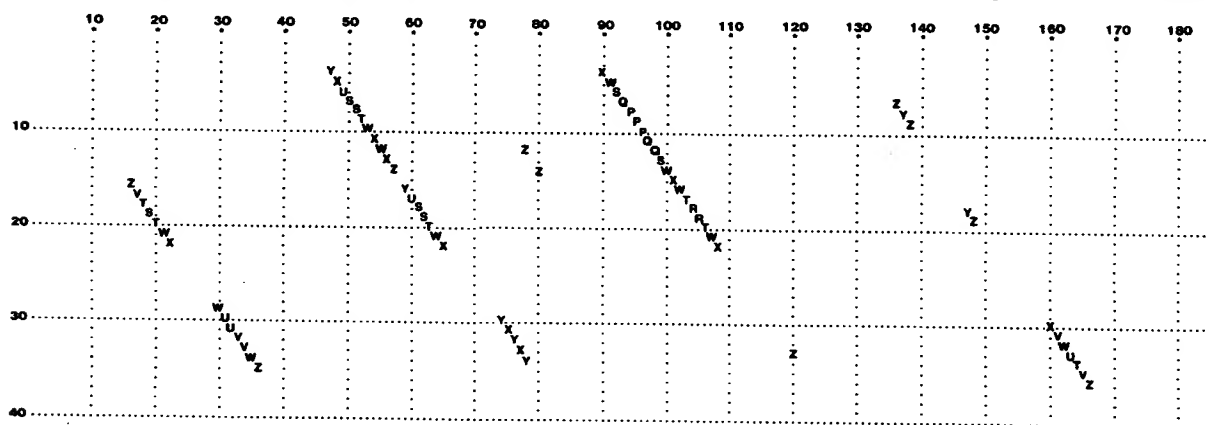

FIG. 5.   Amino-acid homology matrix of WGA (x axis) and chitinase from *P. vulgaris* (y axis). Homology matrices were plotted with the Pustell and Kafatos sequence analysis program (26) using the following parameters: range = 10, scale factor = 0.75, minimum value = 50, compression = 1. Each letter represents homology at that point in the matrix where A = 100%, B = 98%, . . ., Z = 50%. Only homologies in the first 40 amino acids of chitinase are plotted; the remainder of the protein shows no homology with WGA.

Botany: Raikhel and Wilkins

Proc. Natl. Acad. Sci. USA 84 (1987)   6749

four positions (37, 109, 123, and 171) between the deduced amino acids and isolectin 2 could be because of inaccuracies resulting from cross-contamination of the isolectins during fractionation.

Abscisic acid treatment of developing wheat embryos has been shown to affect temporal expression of WGA (12, 15). Using clone pNVR1 as a hybridization probe, we found that abscisic acid treatment of excised wheat embryos modulates mRNA levels for WGA, which is consistent with known effects of abscisic acid on lectin levels (15). Similar results were reported by Williamson et al. (21) for the abundant embryo storage protein. It is possible that abscisic acid regulation is based upon changes in the rates of mRNA transcription, turnover, or processing. It also needs to be mentioned that clone pNVR1 may be hybridizing to the mRNAs for isolectins 1 and 2, as well as to the mRNA for isolectin 3 on the RNA blot. Given the similarity of the isolectin sequences, the high-stringency conditions used for hybridization may not have prevented cross-hybridization with mRNAs from related isolectins.

WGA and chitinase are two chitin-binding proteins that are thought to have antimicrobial activity (34). Recently, however, evidence was presented to show that antifungal activity of WGA can result from contamination by chitinase (35). Comparison of amino acid sequences demonstrated a striking homology between the amino terminus of chitinase (16) and four regions of the WGA molecule. The amino acid residues of WGA directly involved in primary sugar-binding sites are tyrosine-73, serine-62, and glutamic acid-115 (36). These three residues are found in the regions of homology between chitinase and WGA. One may speculate that these regions of homology account for the similarity in chitin-binding activity of these proteins and, subsequently, in copurification. Additionally, the sequence homology between WGA and chitinase may be of functional significance.

1. Goldstein, I. J. & Hayes, C. E. (1978) Adv. Carbohydr. Chem. Biochem. 35, 127–340.
2. Lis, H. & Sharon, N. (1981) Annu. Rev. Biochem. 55, 35–67.
3. Allen, A. K., Neuberger, A. & Sharon, N. (1973) Biochem. J. 131, 155–162.
4. Nagata, Y. & Burger, M. M. (1974) J. Biol. Chem. 10, 3116–3122.
5. Rice, R. H. & Etzler, M. E. (1975) Biochemistry 14, 4093–4099.
6. Peumans, W. J., Stinissen, H. M. & Carlier, A. R. (1982) Planta 154, 562–567.
7. Wright, C. S., Gavilanes, F. & Peterson, D. L. (1984) Biochemistry 23, 280–287.
8. Wright, C. S. & Olafsdottir, S. (1986) J. Biol. Chem. 16, 7191–7195.
9. Raikhel, N. V. & Pratt, L. (1987) Plant Cell Rep. 6, 146–149.
10. Mishkind, M. L., Raikhel, N. V., Palevitz, B. A. & Keegstra, K. (1982) J. Cell Biol. 92, 753–764.
11. Raikhel, N. V., Mishkind, M. L. & Palevitz, B. A. (1984) Planta 162, 55–61.
12. Raikhel, N. V. & Quatrano, R. S. (1986) Planta 168, 433–440.
13. Mishkind, M. L., Palevitz, B. A. & Raikhel, N. V. (1983) Science 220, 1290–1292.
14. Stinissen, H. M., Chrispeels, M. J. & Peumans, W. J. (1985) Planta 164, 278–286.
15. Triplett, B. A. & Quatrano, R. S. (1982) Dev. Biol. 91, 491–496.
16. Broglie, K. E., Gaynor, J. J. & Broglie, R. M. (1986) Proc. Natl. Acad. Sci. USA 83, 6820–6824.
17. Matteucci, M. D. & Caruthers, M. H. (1980) Tetrahedron Lett. 21, 219–722.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
19. Hanahan, D. (1985) in DNA Cloning, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 109–135.
20. Hanahan, D. & Meselson, M. (1983) Methods Enzymol. 100, 333–342.
21. Williamson, J. D., Quatrano, R. S. & Cuming, A. C. (1985) Eur. J. Biochem. 152, 501–507.
22. Thomas, P. S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
23. Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
24. Kyte, J. & Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132.
25. Wright, C. (1987) J. Mol. Biol. 194, 501–529.
26. Pustell, J. & Kafatos, F. C. (1984) Nucleic Acids Res. 12, 643–655.
27. Mansfield, M. A., Peumans, W. & Raikhel, N. V., Planta, in press.
28. Chrispeels, M. J. (1984) Philos. Trans. R. Soc. London Ser. B 304, 309–322.
29. Raikhel, N. V., Mishkind, M. & Palevitz, B. A. (1984) Protoplasma 121, 25–33.
30. Ericson, M. L., Rodin, J., Lenman, M., Glimelius, K., Josefsson, L. G. & Rask, L. (1986) J. Biol. Chem. 31, 14576–14581.
31. Herman, E. M., Shannon, L. M. & Chrispeels, M. J. (1985) Planta 165, 23–29.
32. Faye, L. & Chrispeels, M. J. (1987) Planta 179, 217–224.
33. Dean, D., Tamaki, S., Dunsmuir, P., Favreau, M., Katayama, C., Dooner, H. & Bedbrook, J. (1986) Nucleic Acids Res. 14, 2229–2240.
34. Mirelman, D., Galun, E., Sharon, N. & Loyan, R. (1975) Nature (London) 256, 414–416.
35. Schulumbaum, A., Mauch, F., Vogeli, U. & Boller, T. (1986) Nature (London) 324, 356–367.
36. Wright, C. S. (1984) J. Mol. Biol. 178, 91–104.

Different Legumin Protein Domains Act as Vacuolar Targeting Signals

Gerhard Saalbach,[1] Rudolf Jung, Gotthard Kunze, Isolde Saalbach, Klaus Adler, and Klaus Müntz

Institute of Genetics and Crop Plant Research, Corrensstrasse 3, O-4325 Gatersleben, Sachsen-Anhalt, Federal Republic of Germany

Legumin subunits are synthesized as precursor polypeptides and are transported into protein storage vacuoles in field bean cotyledons. We expressed a legumin subunit in yeast and found that in these cells it is also transported into the vacuoles. To elucidate vacuolar targeting information, we constructed gene fusions of different legumin propolypeptide segments with either yeast invertase or chloramphenicol acetyltransferase as reporters for analysis in yeast or plant cells, respectively. In yeast, increasing the length of the amino-terminal segment increased the portion of invertase directed to the vacuole. Only the complete legumin α chain (281 amino acids) directed over 90% to the vacuole. A short carboxy-terminal legumin segment (76 amino acids) fused to the carboxy terminus of invertase also efficiently targeted this fusion product to yeast vacuoles. With amino-terminal legumin-chloramphenicol acetyltransferase fusions expressed in tobacco seeds, efficient vacuolar targeting was obtained only with the complete α chain. We conclude that legumin contains multiple targeting information, probably formed by higher structures of relatively long peptide sequences.

INTRODUCTION

The secretory system is used to transport proteins via the endoplasmic reticulum (ER) and the Golgi apparatus to the cell surface or to the vacuole/lysosome (Schekman, 1985; Jones and Robinson, 1989; Chrispeels, 1991). Translocation across the ER membrane is mediated by an amino-terminal signal sequence (Perara and Lingappa, 1988). Secretion occurs by default, whereas positive sorting information is required for transport to the vacuole (Kelly, 1985; Pfeffer and Rothman, 1987; Wieland et al., 1987; Dorel et al., 1989).

In mammalian cells, transport to the lysosomes is mediated by mannose-6-phosphate (von Figura and Hasilik, 1986), but in yeast and plant systems, glycans do not act as vacuolar sorting signals (Schwaiger et al., 1982; Voelker et al., 1989). In the yeast hydrolases carboxypeptidase Y (CPY) (Johnson et al., 1987; Valls et al., 1987) and proteinase A (Klionsky et al., 1988), vacuolar sorting information resides in short amino acid sequences at the amino terminus. In plants, Tague et al. (1990) found that the 43 amino-terminal amino acids of mature phytohemagglutinin (PHA) from the common bean are also sufficient to sort invertase to the yeast vacuole. A set of four contiguous amino acids (QRPL) in the CPY signal was identified as critical for vacuolar localization (Valls et al., 1990). For PHA, the sequence LQR was also found to be important (Tague et al., 1990). However, the proteinase A sorting

signal does not resemble the CPY targeting element and the PHA LQR sequence is only partially conserved among lectins. Changes in this region in native PHA do not result in higher levels of secretion, indicating the presence of multiple targeting information in PHA.

The study of PHA-invertase fusions in *Arabidopsis* showed that the PHA segments that target invertase to yeast vacuoles are not sufficient for transport to plant vacuoles, indicating that differences exist between yeast and plant vacuolar sorting processes (Chrispeels, 1991). In a different lectin from barley, a carboxy-terminal propeptide domain of 15 amino acids is necessary for targeting to the vacuoles of tobacco cells (Bednarek et al., 1990).

Here we report vacuolar sorting results obtained from the analysis of the 11S globulin legumin, the major field bean seed protein that accumulates in cotyledon cell storage vacuoles. This protein contains six similar subunits, each composed of two disulfide-linked chains arising from a common precursor (Bassüner et al., 1983; Horstmann, 1983). An amino-terminal signal sequence mediates the cotranslational insertion into the lumen of the ER (Bassüner et al., 1984) and the propolypeptide is then transported by way of the Golgi apparatus (Zur Nieden et al., 1984) into the storage vacuoles. In the ER, 11S propolypeptides are assembled into trimers, whereas hexamers are formed only in the storage vacuoles (for review, see Akazawa and Hara-Nishimura, 1985). The 11S globulins do not contain any sites for *N*-linked glycosylation (Bäumlein et al., 1986;

---

[1] To whom correspondence should be addressed.

Schlesier et al., 1990) and no other form of glycosylation could be detected (Croy et al., 1979; Hurkman and Beevers, 1980).

We expressed in yeast a native legumin propolypeptide gene and gene fusions between different amino-terminal and carboxy-terminal segments of that gene (Baumlein et al., 1986) and the *SUC2* yeast invertase gene (Taussig and Carlson, 1983). We demonstrated that legumin is transported to the yeast vacuoles and that vacuolar targeting information resides both in long amino-terminal and short carboxy-terminal segments. In addition, fusions of amino-terminal legumin segments with chloramphenicol acetyltransferase (CAT) were expressed in tobacco seeds. In plant cells, long amino-terminal segments are also necessary for targeting to vacuoles. However, our results also show that multiple targeting information occurs in the legumin protein and that this information is contained in complex structures rather than in distinct short amino acid sequences.

RESULTS

Expression of Legumin in Yeast

Restriction fragments of a genomic legumin clone LeB4 (Baumlein et al., 1986) and a corresponding cDNA clone B273 (G. Saalbach, unpublished results) encoding a legumin subunit of the B type (Horstmann, 1983) were combined to obtain a complete legumin gene without introns. A SalI restriction site was generated by in vitro mutagenesis 6 bp in front of the ATG start codon, and a SalI-SphI fragment comprising the legumin coding sequence plus 6 bp 5′-untranslated and 143 bp 3′-untranslated sequence was inserted into the yeast plasmid pAAH5 (Ammerer, 1983) where the gene is under the control of alcohol dehydrogenase 1 (*ADH1*) regulatory sequences.

In addition to the original legumin gene, a legumin gene modified by a frameshift starting 36 codons upstream of the stop codon (Saalbach et al., 1988) was used. In the corresponding polypeptide, the 36 carboxy-terminal amino acids would be substituted by a different piece of 52 amino acids.

The plasmids were transformed into yeast, and SDS extracts from these strains were analyzed in protein gel blotting experiments. Figure 1 shows that both the original and the modified proteins could be detected in yeast cells. The intact legumin accumulated to a level of approximately 0.5% of total yeast proteins, but the amount of the modified protein was reduced. Only very low levels of legumin could be detected in the periplasmic space and the culture medium. Estimates from protein gel blots revealed that less than 1% of the expressed legumin was secreted from the yeast cells (data not shown). In yeast, the legumin precursor is not cleaved into the α and β chains.
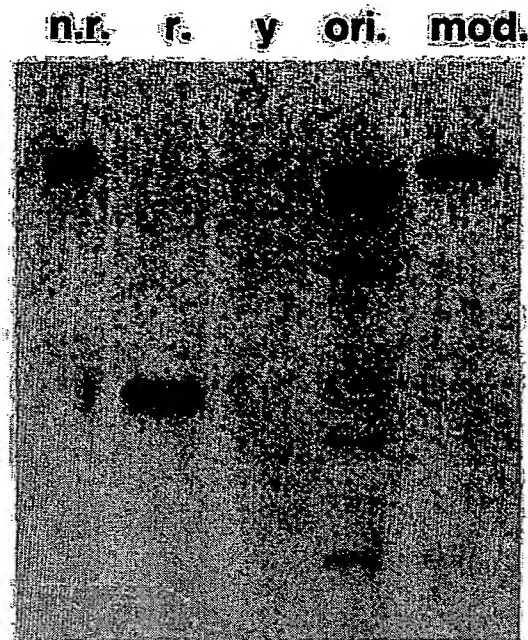
Figure 1. Protein Gel Blot Showing the Expression of Legumin in Yeast.

Lanes 1 and 2, native legumin as marker, nonreduced (n.r.), reduced (r.; the band in this lane represents the α chain; the β chain is not recognized by the antibodies used); lane 3 (y), extract from wild-type yeast as control; lanes 4 and 5, extracts from transformed yeast strains harboring the original (ori.) and the modified (mod.) legumin genes, respectively.

Legumin Is Transported to Yeast Vacuoles

Immunocytochemistry and vacuole isolation procedures were used to detect the intracellular localization of legumin in yeast. Thin sections for electron microscopy were treated with legumin antibodies (gift of R. Manteuffel, Gatersleben) and protein A-gold. The very clear labeling of vacuoles, as shown in Figure 2, indicates efficient targeting of legumin to yeast vacuoles.

The modified legumin could be detected in vacuoles isolated on a discontinuous Ficoll gradient (Stevens et al., 1982). The vacuole fraction and the corresponding total lysate were assayed for α-mannosidase activity (vacuolar marker) and both samples were adjusted to equal α-mannosidase concentration. The modified legumin was detected in these samples by protein gel blotting. Equal legumin band intensities were observed on the blot shown in Figure 3, allowing the conclusion that this modified legumin was also efficiently transported to the yeast vacuole.
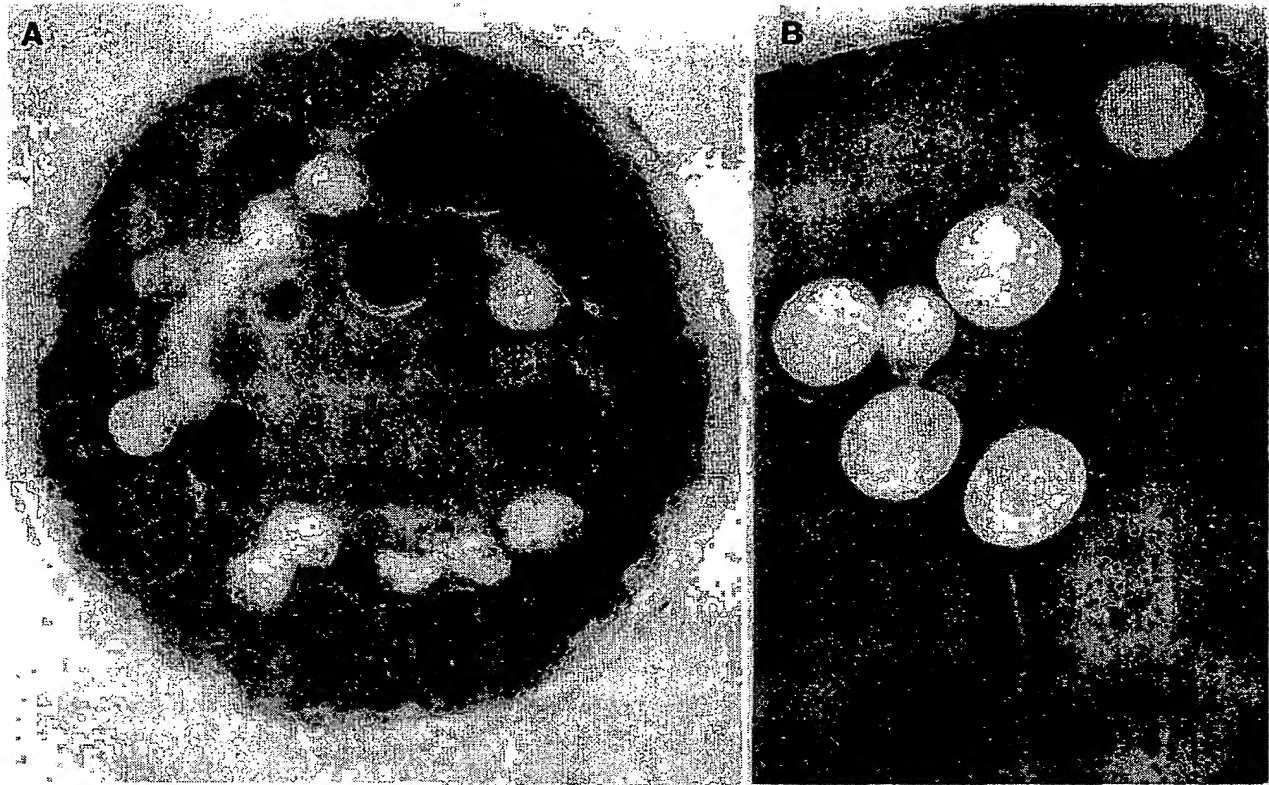
Figure 2. Immunocytochemical Localization of Legumin in Vacuoles of Transgenic Yeast.

Gold particles label the vacuoles very clearly and are completely absent from all other cell compartments.
(A) Magnification ×19,200.
(B) Magnification ×24,400.

Construction of Legumin-Invertase, Invertase-Legumin, and Legumin-CAT Fusion Genes

To determine which sequences in the legumin subunit contain vacuolar targeting information, a series of gene fusions of both amino-terminal and carboxy-terminal segments of legumin with yeast invertase or bacterial CAT were constructed, as shown in Figure 4. For the amino-terminal fusions, the intron-free legumin gene was cloned in front of a SUC2 (invertase) or a CAT gene to create legumin-invertase or legumin-CAT fusions. Carboxy-terminal portions of legumin were removed by deletion mutagenesis using oligonucleotides 24 to 36 bases long comprising the junction sequences. For the carboxy-terminal fusions, a restriction fragment encoding the carboxy-terminal 76 amino acids of legumin was cloned behind the SUC2 gene. Sequences from the amino terminus of the legumin fragment were removed by deletion mutagenesis.

By using oligonucleotide-directed mutagenesis, every desired fusion could be generated exactly. All fusion sequences were verified by sequence analysis.

Seven carboxy-terminal legumin segments ranging from 13 to 76 amino acids in length were fused to the complete invertase, carrying its own signal sequence (Figure 4C). In the invertase-legumin (Inv-Le-C76-38) fusion, 38 amino acids were deleted from the carboxy-terminal end of Inv-Le-C76.

For expression in yeast, appropriate restriction fragments with only 6-bp untranslated legumin 5' sequence (see above) or 26-bp untranslated SUC2 5' sequences plus 15-bp linker were inserted into pAAH5, as described above, and transformed into the strain SEY6210 (suc2-Δ9).

Legumin-CAT fusions were expressed under the control of the strong seed-specific legumin promoter (Bäumlein et al., 1987, 1988) in transgenic tobacco. In addition, several
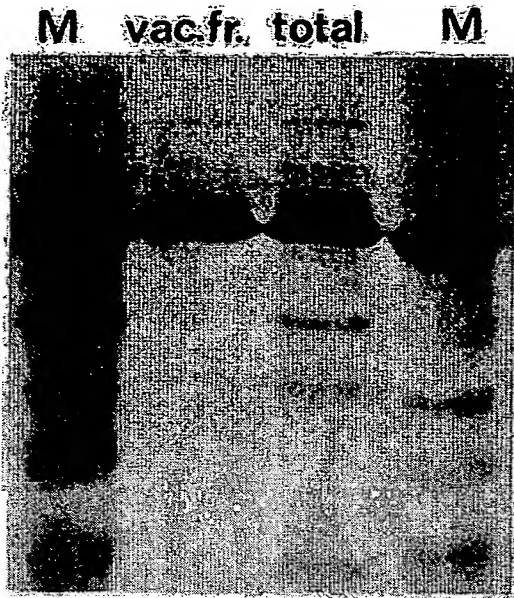
Figure 3. Detection by Immunoblotting of the Modified Legumin in Isolated Yeast Vacuoles.

Left lane (M), extract from yeast cells bearing the modified legumin; right lane (M), extract from untransformed yeast; two middle lanes, samples prepared from isolated vacuoles (vac.fr.) (see Methods) and total spheroplast lysate (total) of yeast cells bearing the modified legumin.

of these fusions were also expressed with the cauliflower mosaic virus (CaMV) 35S promoter in tobacco, and the Le-CAT-462 fusion was expressed in yeast, as described above, to compare the fate of the fusion proteins in the unicellular eukaryote and in different tissues of a higher plant.

Long Amino-Terminal Segments of Legumin Are Necessary To Direct Legumin-Invertase Fusions to Yeast Vacuoles

Results from yeast strains bearing amino-terminal legumin-invertase fusions are summarized in Table 1. The legumin amino acid sequence is shown in Figure 5. We found that efficient secretion occurred when the invertase signal sequence was replaced exactly by the legumin signal sequence. With up to 28 additional amino acids, absolutely no reduction of secretion was observed. Up to 62 additional amino acids retained only low and variable amounts of invertase inside the cells. The segment with an additional 86 amino acids retained 33% of the protein inside the cells. The same result (33% inside the cells) was obtained with
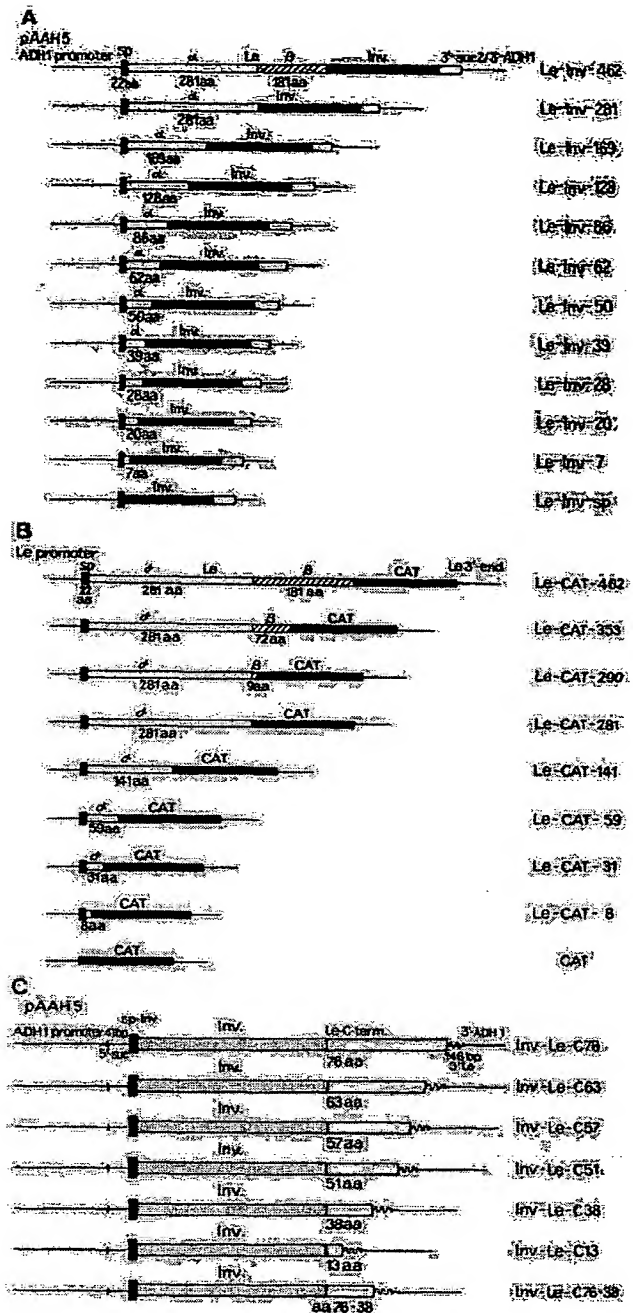
Figure 4. Schematic Representation of the Gene Fusion Constructs Used for the Targeting Experiments in Yeast and Tobacco.

(A) Fusions between amino-terminal legumin segments and yeast invertase.

(B) Fusions between amino-terminal legumin segments and CAT.

(C) Fusions between invertase and carboxy-terminal legumin segments.

Table 1. Invertase Activity Produced by Amino-Terminal Legumin-Invertase Fusions in Yeast

	Invertase Activity[a]		
Fusion	Total	External	% Secretion
Le-Inv-sp	350	350	100
Le-Inv-7	2500	2500	100
Le-Inv-20	1870	1870	100
Le-Inv-28	2000	2000	100
Le-Inv-39	1100	1040	94
Le-Inv-50	670	570	85
Le-Inv-62	1800	1705	95
Le-Inv-86	320	214	66
Le-Inv-128	330	220	66
Le-Inv-169	705	350	50
Le-Inv-281	600	45	7
Le-Inv-462	–[b]		

[a] Invertase activity is in units (nanomoles of glucose per minute at 30°C) per OD_{600} yeast cells.
[b] Fusion unstable in pAAH5.

128 amino acids. A segment with 169 amino acids kept 50% intracellular, and the whole α chain (281 amino acids) was necessary to retain more than 90% inside the yeast cells.

Native legumin is transported to the yeast vacuoles. To determine whether the legumin segments also target the intracellular portions of invertase activity to the vacuoles, we isolated vacuoles from the strain harboring the Le-Inv-86 fusion, retaining 33% of the activity inside the cell. Recovery of vacuoles was calculated from the α-mannosidase (vacuolar marker) activity in the spheroplast lysate and in the vacuole fraction. Contamination with ER or cytoplasma was low (approximately 10%), as determined by assaying the fractions for cytochrome c reductase and α-glucosidase, respectively. The recovery of intracellular invertase in the vacuole fraction was calculated from the activities in the whole cells, in the spheroplast supernatant, in the spheroplast lysate, and in the vacuole fraction. Figure 6 shows that the intracellular invertase activity cofractionated with the vacuolar marker α-mannosidase, indicating that the legumin fusions are targeted to the vacuoles.

A Short Carboxy-Terminal Legumin Segment Efficiently Targets Invertase to Yeast Vacuoles

The carboxy-terminal invertase-legumin fusions were analyzed in yeast in the same way as described for the amino-terminal fusions. The results are shown in Table 2. A short segment of 13 amino acids caused practically no reduction of secretion. In the case of the Inv-Le-C38 fusion, the total cellular invertase activity was very low, indicating a structural interaction between a specific legumin segment and

invertase, and 28% of that activity was retained inside the cells. With an additional 13 amino acids (Inv-Le-C51), activity was normal and only 17% was retained. Six additional amino acids (Inv-Le-C57) caused an increase in the retention to 43% inside the cells and another 6 amino acids (Inv-Le-C63) increased that value to only 54%. With the addition of the last 76 carboxy-terminal amino acids of legumin (for the sequence, see Figure 5) to invertase, practically complete (93%) retention was achieved. This activity was also transported efficiently to the yeast vacuoles, as shown by vacuole isolation described above (Figure 6). These results indicated that the essential sequence in the carboxy-terminal signal might be located in the amino-terminal part of the 76-amino acid segment. However, deletion of the carboxy-terminal half (Inv-Le-C76-38) resulted in considerable loss of vacuolar transport to only 50%, indicating that the complete segment is necessary to form the signal.

Does the Level of Expression Influence the Secretion/Retention Ratio?

The total invertase activities expressed from the different gene fusions ranged from about 200 units per OD of cells (Inv-Le-C76) to 2500 units per OD of cells (Le-Inv-7). These differences could be due to the influence of the legumin segments on the specific activities of the fusion molecules. In case of CPY-invertase and PHA-invertase fusions, activities were not as variable. Values of approximately 200 to 300 and 350 to 750 units per OD of cells, respectively, were observed (Johnson et al., 1987; Tague et al., 1990). About threefold enhancement of the expression of a PHA-Inv fusion caused an increase in secretion from 10% to 22%, probably because of the saturation of the vacuolar sorting machinery (Tague et al., 1990). To analyze whether

```
       TSSEFDRLNQ CRLDNINALE PDHRVESEAG LTETWNPNHP  40
       ELRCAGVSLI RRTIDPNGLH LPSYSPSPQL IYIIQGKGVI  80
       GLTLPGCPQT YQEPRSSQSR QGSRQQQPDS HQKIRRFRKG  120
       DIIAIPSGIP YWTYNNGDEP LVAISLLDTS NIANQLDSTP  160
       RVFYLGGNPE VEFPETQEEQ QERHQQKHSL PVGRRGGQHQ  200
       QEEESEEQKD GNSVLSGFSS EFLAQTFNTE EDTAKRLRSP  240
       RDKRNQIVRV EGGLRIINPE GQQEEEEQEE EEKQRSEQGRN 281

181    GLEETICSLK IRENIAQPAR ADLYNPRAGS ISTANSLTLP
141    ILRYLRLSAE YVRLYRNGIY APHWNINANS LLYVIRGEGR
101    VRIVNSQGNA VFDNKVRKGQ LVVVPQNFVV AEQAGEEEGL
 61    EYLVFKTNDR AAVSHVQQVF RATPADVLAN AFGLRQRQVT
 21    ELKLSGNRGP LVHPQSQSQS N
```

Figure 5. Amino Acid Sequence of the Legumin Propolypeptide.

The sequence was derived from a genomic DNA sequence (Bäumlein et al., 1986) and is shown without the signal peptide. The upper part numbered from 1 to 281 (starting at the amino terminus) is the sequence of the α chain. The lower part numbered from 1 to 181 (starting at the carboxy terminus) is the sequence of the β chain.
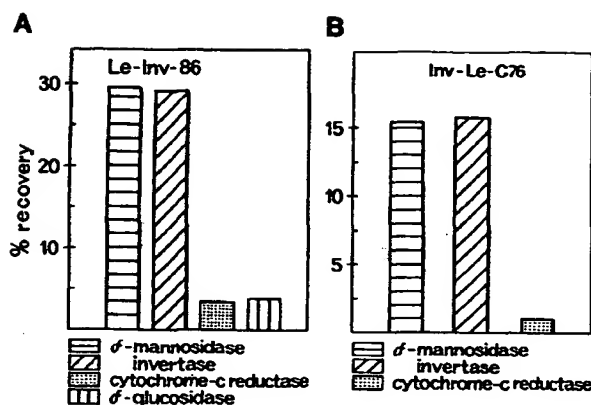
A

Le-Inv-86

% recovery

30

20

10

☰ *d*-mannosidase
▨ invertase
▦ cytochrome-c reductase
▥ *d*-glucosidase

B

Inv-Le-C76

15

10

5

☰ *d*-mannosidase
▨ invertase
▦ cytochrome-c reductase

Figure 6. Localization of Legumin-Invertase Fusions in Isolated Yeast Vacuoles.

Vacuoles were isolated (see Methods) and the indicated marker enzyme activities were determined in the total lysate and in the vacuole fraction.

(A) and **(B)** Cofractionation of invertase activity of fusions Le-Inv-86 **(A)** and Inv-Le-C76 **(B)** with the vacuolar marker enzyme *a*-mannosidase.

the level of expression of our legumin-invertase fusions allowed correct results, we also expressed the fusion Le-Inv-86 at a much lower level by transferring the fusion including *ADH1* promoter and terminator from the multicopy plasmid pAAH5 to the yeast plasmid YCp50 (Sherman et al., 1986). After transformation with this plasmid, only one copy of the fusion gene was present per yeast cell, resulting in a very low expression. With this low activity, variable results were obtained because of background interference ranging from 30 to 70% (average 50%) secretion (data not shown). This result indicated a certain overloading of the vacuolar transport pathway. On the other hand, retention values went up to more than 90% with long legumin segments. In addition, through the use of immunocytochemistry, native legumin expressed with pAAH5 could only be detected in yeast vacuoles, and only trace amounts of legumin were detected in periplasma and medium samples (see above). Together, these results indicated that the conclusions derived from the use of the yeast expression system are essentially correct.

Glycosylation of Invertase Provides Additional Evidence for Passage through the Secretory System

Secretory invertase is highly and heterogeneously glycosylated. This glycosylation occurs in the ER and in the Golgi apparatus. Wild-type yeast cells also produce a

cytoplasmic and, therefore, unglycosylated form of invertase (Carlson and Botstein, 1982; Perlman et al., 1982). These two forms can be distinguished easily by their different mobility in a native acrylamide gel (Gabriel and Wang, 1969; Carlson et al., 1981).

All amino-terminal and carboxy-terminal legumin-invertase fusions (Figures 4A and 4C) produced in yeast were analyzed for invertase glycosylation. In all cases, the invertase was highly glycosylated, as illustrated in Figure 7 for some examples. This means that all fusions enter the secretory system and pass the Golgi apparatus, yielding additional evidence for the localization of all intracellularly retained fusions in the vacuole.

Analysis of the Expression of Legumin-CAT Fusions in Plants

Although the legumin propolypeptide (this paper) and PHA (Tague et al., 1990) are sorted into the vacuoles of yeast, it remains to be demonstrated that these cells use the same sorting mechanism and recognize the same targeting signals as the storage tissue cells in developing plant seeds. Field bean, the donor of the legumin gene, still cannot be transformed and regenerated efficiently. Therefore, we have used the tobacco transformation system to verify our results in plants. Like many other seed storage protein genes, the legumin gene used in our study is correctly expressed in tobacco seeds (Bäumlein et al., 1987). It has also been shown that plant vacuolar proteins such as phaseolin and PHA from common bean and the vegetative storage protein patatin from potato are transported to the vacuoles in transgenic tobacco (Greenwood and Chrispeels, 1985; Sturm et al., 1988; Sonnewald et al., 1989).

Legumin-CAT gene fusions (Figure 4B) were transformed into tobacco using the Ti plasmid system. Transformation and expression of the gene fusions were verified by DNA and RNA blotting techniques. The RNA gel blot shown in Figure 8 reveals that mRNA is formed from all

Table 2. Invertase Activity Produced by Carboxy-Terminal Legumin-Invertase Fusions in Yeast

Fusion	Invertase Activity		
	Total	External	% Secretion
Inv-Le-C13	900	845	94
Inv-Le-C38	5	3.6	72
Inv-Le-C51	500	415	83
Inv-Le-C57	440	250	57
Inv-Le-C63	550	250	46
Inv-Le-C76	190	13	7
Inv-Le-C76-38	580	290	50
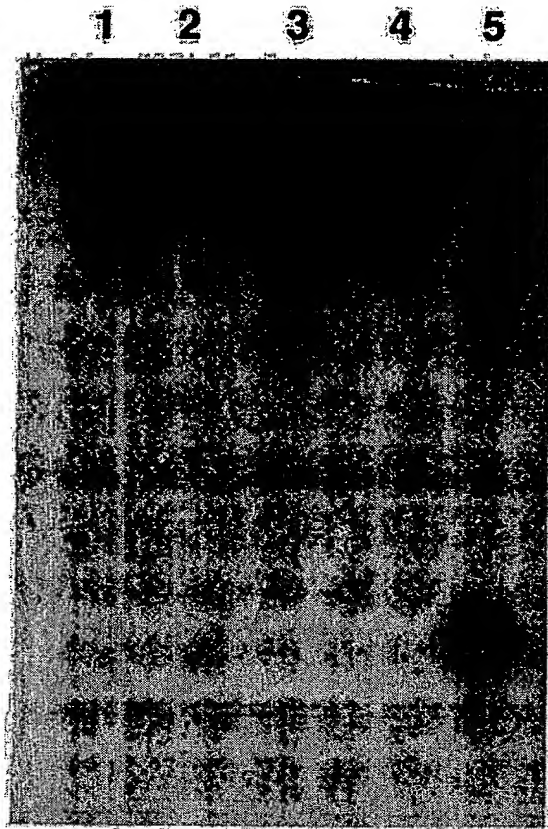
Figure 7. Invertase Activity Gel Demonstrating the Glycosylation of Legumin-Invertase Fusions in Yeast.

Lane 1, wild-type invertase (purchased sample, Boehringer Mannheim) as marker; other lanes show invertase activity of fusions Le-Inv-sp (lane 2), Le-Inv-86 (lane 3), Inv-Le-C76 (lane 4). All invertase fusion proteins are highly glycosylated, as can be seen from comparison with a cytoplasmic unglycosylated invertase activity in lane 5 resulting from a construct with incomplete signal sequence.

transformed fusion genes at about the same steady-state level in seeds when controlled by the legumin promoter and in seeds and leaves when programmed by the CaMV 35S promoter.

Protein level expression was analyzed by the CAT activity assay and by immunoblotting with anti-legumin and anti-CAT antibodies. CAT activity could only be detected in seeds with the fusions Le-CAT-59, Le-CAT-8, and CAT. Four of the fusions without CAT activity were expressed in leaves with the CaMV 35S promoter. Also in this case no CAT activity could be detected. That the long legumin-CAT fusions can in principle exhibit CAT activity was demonstrated by expression of the Le-CAT-462 fusion in

yeast. As shown in Figure 9, high CAT activity could be observed.

The results from the CAT activity assays were confirmed by immunoblotting using a CAT antiserum. CAT protein could only be detected in seeds transformed with Le-CAT-59, Le-CAT-8, and CAT (data not shown). Figure 10 shows that by using anti-legumin antibodies, the α chain and a degradation product of legumin could be detected in seeds for fusion constructs Le-CAT-290, Le-CAT-353, and Le-CAT-462. No legumin could be detected with Le-CAT-281. In this fusion, the β chain was removed and substituted exactly by CAT, probably destroying the recognition and cleavage sequences for the α-β cleavage. These results indicated that as long as the proteolytic cleavage sequence is preserved in the fusions (which is obviously the case also with only 9 amino acids of the β chain in the fusion Le-CAT-290), proteolytic cleavage occurs at this site.

It is suggested, therefore, that these fusions are transported into the protein bodies because it is known that specific processing of the propolypeptides of 11S globulin subunits takes place in the protein storage vacuoles (Chrispeels et al., 1982; Hara-Nishimura and Hara-Nishimura, 1987). The β chain-CAT portion is obviously degraded because no CAT activity is present and β chain-specific antibodies do not detect any additional band (data not shown). The occurrence of a degradation product larger in size than the β chain indicates that the α chain is also partially degraded (Figure 10). The shorter the β chain portion in the fusions, the greater is the degradation. In
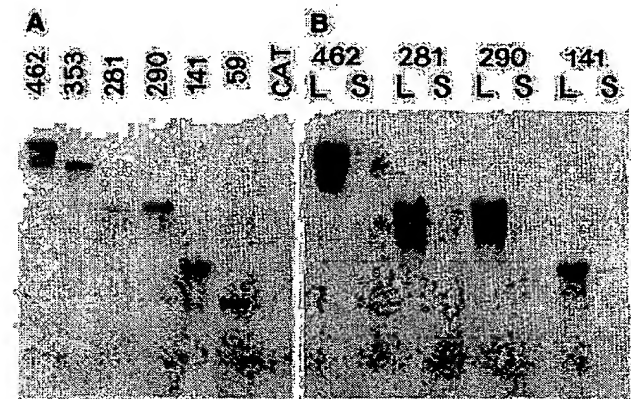
Figure 8. RNA Gel Blot Showing the Expression of Legumin-CAT Fusion Genes in Transgenic Tobacco.

Le-CAT fusions are indicated by the number of legumin amino acids.
(A) Expression of the legumin promoter in seeds.
(B) Expression of the CaMV 35S promoter in leaves (L) and seeds (S).

the leaves of the CaMV 35S promoter transformants, no legumin could be detected by immunoblotting, indicating complete degradation of the fusions in the leaf cells where these long fusions are probably also transported to the vacuoles (see Discussion).

Vacuolar Targeting of Legumin-CAT Fusions in Tobacco Seeds

Several fusions were analyzed for their intracellular localization. For this purpose, protein bodies were isolated from tobacco seeds using a potassium iodide step gradient in glycerol (Sturm et al., 1988). The fractions were analyzed either for CAT activity (in the case of the CAT fusions Le-CAT-8 and Le-CAT-59) or for the presence of the legumin α chain and its degradation product (in the case of the fusions Le-CAT-290 and Le-CAT-462). As shown in Figure 9, the CAT activity produced by the three short legumin-CAT fusions could not be found in the protein body fraction, whereas the α chain (in the case of Le-CAT-462) or the degradation product (in the case of Le-CAT-290) was present in the isolated protein bodies, as demonstrated in Figure 10. These results showed that in plants, as in yeast, short amino-terminal segments of legumin are not sufficient for vacuolar targeting, whereas segments as long as the α chain are able to direct the fusions to the vacuoles.

DISCUSSION

Vacuolar Targeting Information in the Legumin Propolypeptide

The aim of this work was to analyze the targeting information of the plant vacuolar protein legumin, the major field bean seed storage protein. We used both yeast and plants for our intracellular transport analysis. Yeast has been used for the investigation of vacuolar targeting information of yeast proteases (Johnson et al., 1987; Valls et al., 1987, 1990; Klionsky et al., 1988) and has been shown to be suitable for the analysis of the plant vacuolar protein PHA (Tague and Chrispeels, 1987; Tague et al., 1990; Chrispeels, 1991).

We expressed native legumin in yeast and found that the legumin subunit precursor accumulates inside the yeast cells. We also demonstrated that this plant vacuolar protein is efficiently transported to yeast vacuoles. To localize the vacuolar targeting information in the legumin propolypeptide, a series of gene fusions between segments of legumin and yeast invertase were constructed and analyzed in yeast. We found that in contrast to the results obtained with yeast proteases and PHA, short amino-terminal segments of legumin up to about 40 amino acids additional to the signal peptide were not able to

Figure 9. CAT Activity Assay of Cell Fractions from Transgenic Tobacco Seeds.

Left two lanes, comparison of expression of Le-CAT-462 in yeast and tobacco. The fusion with complete legumin yields CAT activity in yeast but not in tobacco seeds. Right three sections, detection of CAT activity in different fractions from the potassium iodide step gradient. Lanes 4 indicate the protein body fraction prepared from seeds expressing the indicated short legumin-CAT fusions; no CAT activity was found in the protein body fractions.

retain any portion of the fusions inside the cells. The vacuolar portion increased with increasing length of the legumin segment. A segment of 86 additional amino acids targets 33% of the invertase activity to the vacuole and only the whole α chain (281 amino acids) targets more than 90% to the vacuole.

The results obtained with the amino-terminal fusions in yeast were principally confirmed in a plant system. Legumin-CAT fusions were expressed in tobacco leaves and seeds. If the legumin signal peptide plus an additional 8 or 59 amino acids were fused to CAT, it did not reach the vacuoles. The whole legumin (462 amino acids) or the α chain plus 9 amino acids of the β chain allowed for targeting of CAT fusions to the vacuoles, indicating that, as in yeast, short amino-terminal segments of legumin are not sufficient for vacuolar targeting, whereas the α chain is able to direct the fusions to the vacuoles.

The analysis of fusions with carboxy-terminal legumin segments indicated that this region also contains vacuolar targeting information. As in the case of the amino-terminal segments, vacuolar transport increased with increasing length of the carboxy-terminal segments. However, a much shorter segment (76 amino acids) was sufficient for efficient targeting. We conclude from these results that there are at least two regions in the legumin propolypeptide containing vacuolar targeting information and that these regions can act independently of each other.

On the Nature of the Vacuolar Targeting Signal

Interaction of a specific signal and a corresponding receptor is likely for vacuolar targeting of the yeast CPY protein (Valls et al., 1990). The first 30 amino-terminal amino acids
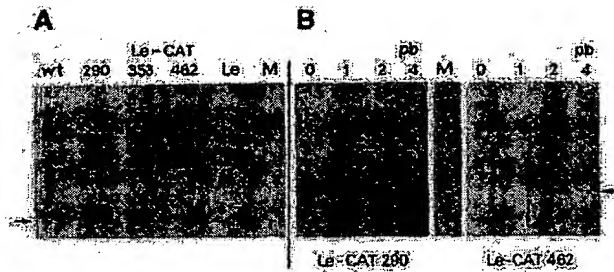
Figure 10. Detection by Immunoblotting of the Legumin Part of Le-CAT Fusions in Transgenic Tobacco Seeds.

(A) Detection of the α chain (upper band) and a degradation product, indicated by the arrow, in seed extracts.
(B) Detection of the α chain (upper band) and a degradation product (arrow) in protein body fractions (pb) isolated from seeds expressing the indicated fusions.
wt, untransformed tobacco; Le, original legumin expressed in tobacco; M, native legumin.

of the CPY propeptide are sufficient and necessary for efficient targeting of CPY-invertase fusions to the vacuole (Johnson et al., 1987). The tetrapeptide QRPL located within this small domain (positions 4 to 8 of proCPY) was found to be critical for targeting (Valls et al., 1990). However, the vacuolar sorting signal of yeast proteinase A (Klionsky et al., 1988) is not similar to the CPY sorting sequence and an element comparable to the QRPL tetrapeptide is not present. This indicates that either different signals exist or the character of the signal is formed by physicochemical and/or structural properties of the targeting element and is not dependent on a highly conserved amino acid sequence.

An amino-terminal segment of about 30 residues of the plant vacuolar protein PHA also contains information for targeting to yeast vacuoles. The sequence resembles that of CPY, and the tetrapeptide LQRD, partially homologous to the QRPL of CPY at positions 18 to 21 of mature PHA, forms the essential core of the signal (Tague et al., 1990). However, this amino-terminal segment is sufficient but not necessary for vacuolar targeting of PHA in yeast, indicating that there is more information in the PHA protein for targeting to the vacuole.

We also found multiple targeting information in the plant vacuolar protein legumin. However, with respect to the nature of the signal, our results differed from those reported by others. In both the amino-terminal and the carboxy-terminal regions, long and complete sequences are necessary for efficient targeting, indicating that the signal is not formed by the primary structure of a short polypeptide sequence. Long polypeptide sequences could form the targeting signal by establishing a "signal patch" on the surface of the protein structure after folding.

On the other hand, this result could also mean that a different sorting mechanism might be involved. In animal cells, secreted proteins (zymogens) are sorted in the Golgi (and in the ER if a certain concentration is reached) by forming protein aggregate granules that specifically exclude other luminal ER proteins (i.e., condensation sorting; see Burgess and Kelly, 1987; Tooze et al., 1989). Such an interaction between the sorted proteins with each other and with a membrane component that is not necessarily a high-affinity receptor could depend on the length of the polypeptide segment. One has also to consider that in the plant, legumin propolypeptides form trimers in the ER and that these trimers are transported from the ER to the Golgi apparatus, where they are sorted in the vacuoles. It is very unlikely that comparable structures are formed by the legumin-invertase hybrid proteins in yeast. It remains to be demonstrated whether the structures found to act as sorting signals in yeast act by way of the same mechanism in plant cells and, if so, how these structures are exposed by the trimeric conformation.

Another possible explanation for the results obtained with plant vacuolar proteins and fusion proteins in yeast is that plant proteins and hybrid proteins are not correctly folded by the protein transport and folding machinery of polypeptide chain-binding proteins (for review, see Rothman, 1989) in yeast. A so-called scavenger pathway might exist, as suggested by Johnson et al. (1987), recognizing misfolded proteins and directing them to the vacuole for degradation. The efficiency of such a mechanism obviously would be highly variable and would depend on the degree of misfolding. On the other hand, our legumin-CAT fusion results in tobacco indicated that the long segments truly might act as transport signals. Preliminary data from the analysis of the legumin-invertase fusions in tobacco mesophyll cells suggest that the legumin α chain also efficiently directs invertase to the vacuoles.

There is further evidence that very complex structures might be necessary to form a plant vacuolar targeting signal. An amino-terminal segment of 113 amino acids of a vacuolar proteinase inhibitor from potato that is only 123 amino acids long still allowed complete secretion of invertase from Arabidopsis cells (von Schaewen et al., 1990). The short PHA segments directing invertase to yeast vacuoles also are not sufficient for vacuolar targeting in plant cells (Chrispeels, 1991). In barley lectin, a carboxy-terminal propeptide domain (only 15 amino acids long) was found to be necessary for targeting to plant vacuoles (Bednarek et al., 1990). It has yet to be shown whether this segment is sufficient for targeting or whether it is involved in the formation of a special protein conformation. The domain has the potential to form an amphipathic helix that might interact with other lectin domains (Bednarek et al., 1990).

Despite the observed differences between the legumin sorting signals and those of PHA, for example, we compared the protein sequences for the tripeptide LQR (part

of the tetrapeptide LQRD, see above) because Tague et al. (1990) found this tripeptide also in the carboxy-terminal region of legumin A from pea, which is homologous to the carboxy-terminal part of field bean legumin found to act as a vacuolar sorting signal. However, in field bean legumin, GLRQR occurs at positions 29 to 25 (numbered from the carboxy terminus as in Figure 5) instead of the NLQRN in pea legumin. Finally, these sequences are located in the carboxy-terminal half of the sorting signal, the deletion of which did not lead to complete loss of vacuolar targeting but only to a reduction from more than 90% with the fusion Inv-Le-C76 to 50% with Inv-Le-C76-38.

In addition, we screened 67 sequences from our data bank of plant vacuolar proteins, mostly targeted into storage vacuoles of seeds, for the occurrence of the sequence LQR. The results are shown in Table 3. Even allowing conservative amino acid substitutions (isoleucine and valine for leucine, asparagine for glutamine, lysine for arginine), only about half of the sequences contained this putative signal. For example, in 11S globulins, only the A-type subunits (e.g., legumin A of pea) contained the tripeptide, even multiple at different sites, whereas it did not occur in B-type subunits such as the one used in our studies.

Taken together, no generalizations can be made about vacuolar sorting signals and mechanisms. On the one hand, relatively short amino acid stretches with similar essential core sequences can act as sorting signals, very likely by way of a specific receptor, and the deletion of short propeptide domains leads to secretion (Chrispeels, 1991). On the other hand, we found both in yeast and in plant cells that vacuolar targeting of legumin is mediated by very different, much more complex sequences that might form the signal by higher structures and could act by way of a less specific mechanism.

METHODS

Reagents

DNA-modifying enzymes were obtained from Boehringer Mannheim and from Bethesda Research Laboratories and were used according to the manufacturers' instructions. DNA sequencing reagents were from Pharmacia and Boehringer Mannheim. Radioactive deoxynucleotides and nucleotides were from Amersham. For protein gel blot staining, Promega's ProtoBlot Western Blot AP System and Amersham's streptavidin-biotin system were used. Novozym 234 and Lysing Enzyme (cell wall lysing enzymes from Trichoderma harzianum) were from Calbiochem and Sigma, respectively. o-Dianisidin, N-ethylmaleimide, and Ficoll 400 were obtained from Sigma. NADPH, p-nitrophenyl-α-o-glucopyranoside, p-nitrophenyl-α-o-mannopyranoside, and low melting agarose were from Serva; horseradish peroxidase, glucose oxidase, and invertase from Boehringer Mannheim; and cytochrome c from Biomed (Cracow, Poland). Antibiotics for selection of transformed

Table 3. Occurrence of the Tripeptide LQR in Plant Vacuolar Proteins

Types of Plant Vacuolar Proteins	No. of Sequences		
	With LQR	Without LQR	Total
11S globulins	13	11	24
7S globulins			
Vicilin-like	8	0	8
Convicilin-like	4	0	4
Lectins	4	3	7
Albumins	3	2	5
2S proteins	1	3	4
Prolamins	0	11	11
Vegetative storage proteins	0	4	4
Total	33	34	67

bacteria and transformed plants were obtained mainly from Serva, and ^{14}C-chloramphenicol was from Amersham. Antisera against CAT were from Hoffmann-LaRoche Inc. (Burns and Crowl, 1987) and from 5 Prime-3 Prime Inc.; the legumin antibody used for immunocytochemistry was a gift of R. Manteuffel (Gatersleben). Affi-Gel was from Bio-Rad and Protein A was from Pharmacia.

Strains and Media

Escherichia coli strains were used for plasmid maintenance. E. coli JM101 was used for production of single-stranded DNA with the helper phage M13K07 (Vieira and Messing, 1988). E. coli GJ23 was used as helper strain for conjugation (van Haute et al., 1983). Agrobacterium tumefaciens C5801 (rifampicin) containing the disarmed Ti plasmid pGV3850 (Zambryski et al., 1983) was used for transformation of tobacco (Nicotiana tabacum Petit Havana cv SR1, Horsch et al., 1985). Nutrient broth (Immunpräparate Berlin) was used for growth of E. coli strains; YEB medium (Vervliet et al., 1975) was used for growth of Agrobacterium strains.

Yeast (Saccharomyces cerevisiae) strain SHY2 (ura3, trp1, leu2, his3, can) (Botstein et al., 1979) was used for expression of native legumin; strain SEY6210 (ura3, trp1, leu2, his3, lys2, can, suc2-Δ9) (obtained from S. Emr, Pasadena, CA) was used for expression of legumin-invertase fusions. For growth of legumin-bearing SHY2 strains, minimal medium with 2% glucose was used according to Tanaka et al. (1967). For SEY6210 strains bearing legumin-invertase fusions, the medium contained 2% fructose as carbon source. Cells for spheroplast formation and vacuole isolation were grown in minimal medium according to Wickerham (1946).

Plasmid Constructions

All basic constructions were performed in the phagemid pBS– (Stratagene). The intron-free legumin gene was constructed by substituting a BamHI-KpnI restriction fragment of a legumin cDNA clone, B273, isolated in our laboratory (G. Saalbach, unpublished

results) for the corresponding intron-containing fragment of a legumin genomic clone, LeB4 (Bäumlein et al., 1986). An appropriate restriction fragment from pSV2CAT (Gorman et al., 1982) containing the entire coding sequence of the CAT gene was inserted 3' of the legumin gene and Le-CAT fusions generated by oligonucleotide-directed deletion mutagenesis. Approximately 470 bp of legumin 3'-untranslated region was fused behind the stop codon of CAT as a polyadenylation signal.

For transformation into tobacco, fusions including the legumin promoter were subcloned into the SmaI site of the plant intermediary vector pMLJ1 (gift of L. Herrera-Estrella). A SalI site was generated 6 bp 5' to the ATG translation start of the legumin gene, and corresponding SalI-HindIII fragments of several fusions were inserted into the SmaI site of pRT103 (Töpfer et al., 1987) for expression with the CaMV 35S promoter. HindIII fragments of these constructs were also subcloned into pMLJ1. An SphI fragment comprising the coding region plus 148 bp of 5'-flanking and 147 bp of 3'-flanking regions was isolated from the intron-free legumin gene and inserted into the SphI site of pBS−, resulting in pLeSph.

Plasmid pSUC23 containing a SUC2 gene (Taussig and Carlson, 1983) was obtained from T. Rapoport (Berlin). A HindIII fragment of this gene spanning bases 12 to approximately 2700 was inserted into the HindIII site of the pLeSph such that both genes were in the same orientation. The legumin-invertase gene fusions were generated by deletion mutagenesis. For the fusion of carboxy-terminal legumin segments to invertase, an EcoRI-PstI fragment from pSUC23 spanning bases −26 (plus polylinker) to approximately 2270 of the SUC2 gene was subcloned into pBS− by way of EcoRI-PstI. The resulting plasmid pSU was opened with PstI, blunt ended by T4-DNA polymerase, and then cut with SphI (polylinker). A KpnI(blunt)-SphI fragment spanning bases 1473 to 1850 of the legumin gene with 230 bp of 3'-coding region and 147 bp of 3'-noncoding region was inserted into the opened pSU plasmid such that the SUC2 gene and the legumin gene fragment were in the same orientation.

Oligonucleotide-directed in vitro mutagenesis for generation of restriction sites and deletions was performed essentially according to Zoller and Smith (1983). Oligonucleotides were synthesized on an Applied Biosystems 391 DNA synthesizer (F. Machemehl, Gatersleben). For deletions, oligonucleotides encoding the junction sequences were 30 to 40 bases long. The fusion sequences were verified using dideoxy DNA sequencing (Sanger et al., 1977).

For expression of legumin and legumin-invertase fusions in yeast, SalI-HindIII (polylinker) restriction fragments starting 6 bp in front of the ATG (see above) were isolated from the pBS− plasmids and inserted by way of blunt-end ligation into the HindIII site of the yeast shuttle plasmid pAAH5 (Ammerer, 1983). For the carboxy-terminal invertase-legumin fusions, EcoRI-SphI fragments were used in the same way. For low-level expression in yeast, a BamHI fragment was isolated from the pAAH5 clone containing Le-Inv-86. This fragment comprising the ADH1 promoter, the legumin invertase gene fusion, and the ADH1 terminator was inserted into the BamHI site of YCp50 (Sherman et al., 1986; obtained from S. Emr, Pasadena, CA).

Transformation of Yeast and Tobacco

Yeast was transformed according to Ito et al. (1983), and transformed cells were selected on minimal medium with the required supplements.

Tobacco leaves were used for Agrobacterium-mediated leaf disc infection as described by Horsch et al. (1985). Transformants were selected on 100 μg/mL kanamycin. Transformation of the genes was verified by DNA gel blotting (Southern, 1975), and expression of the genes by RNA gel blotting using the method of Thomas (1983).

Immunological Procedures and Electron Microscopy

Antibodies against a B-type subunit of legumin (gift of C. Horstmann, Gatersleben) and against α and β chains isolated from a denaturing reducing polyacrylamide gel were raised in mice. The antigens were bound to Affi-Gel and used for affinity purification of the antibodies.

Yeast cells were homogenized with dry ice and glass beads in a mortar, and the frozen powder was extracted by boiling in SDS sample buffer. Tobacco seeds were ground in a mortar under liquid nitrogen, and the powder was extracted in 0.1 M phosphate buffer, pH 7.5, with 1 M KCl. Spheroplasts and isolated vacuoles from yeast as well as cell fractions from tobacco seeds were lysed by addition of concentrated SDS sample buffer and boiling. Samples were separated on polyacrylamide gels (Laemmli, 1970) and blotted to nitrocellulose according to Towbin et al. (1979). Legumin bands were visualized by treating the blots with legumin antibodies, followed by biotinylated anti-mouse immunoglobulin G and streptavidin-alkaline phosphatase or by alkaline phosphatase immunoglobulin G conjugates, and staining with nitro blue tetrazolium and 5-bromo-4-chloroindolylphosphate.

For electron microscopy, yeast cells were prefixed with 2% glutaraldehyde in cacodylate buffer, pH 7.2, with 1 M sorbose. After centrifugation, the cells were wrapped in 3% low melting agarose (gel point 26 to 29°C). Pieces of 1 mm³ were quickly frozen in liquid propane (−185°C) and stored under liquid nitrogen. Freeze-substitution was carried out according to Müller et al. (1980). At −35°C, the samples were embedded in Lowikryl K4M resin, which yields a good preservation of cellular structures and low nonspecific antibody binding (Roth et al., 1981; Craig and Goodchild, 1982). Embedding was carried out in gelatin capsules by UV light polymerization under nitrogen for 24 hr. After 2 days of curing, thin sections were cut using glass knives on an LKB-Ultrotome. Sections were collected on Formvar film-coated nickel grids. Nonspecific binding was blocked by treatment with 0.1% BSA in 0.15 M phosphate buffer, pH 7.2, containing 0.5% Tween 20, 0.1% PEG, and 5 mM ammonium chloride for 10 min. The grids were incubated on drops of diluted legumin antibody solution for 15 min and then extensively washed in blocking solution with reduced BSA content (100 μg/mL). Thereafter, the grids were treated with protein A-gold conjugate prepared according to Roth (1983) for 15 min and washed again. Finally, grids were washed in distilled water and contrasted with uranyl acetate (1% in ethanol). After washing with pure ethanol and drying, specimens were evaluated in a TESLA BS500 transmission electron microscope at 60 kV. Controls were run without antibody treatment and with preimmune serum.

Assays

Quantitative invertase assays were performed using the method of Goldstein and Lampen (1975), as described by Johnson et al.

(1987). External activity was measured using intact cells, total activity after lysis of the cells with 0.5% Triton X-100. Invertase activity gels were performed using triphenyltetrazolium chloride according to Gabriel and Wang (1969). CAT activity was assayed according to Gorman et al. (1982).

Spheroplasts of yeast were prepared as described by Maraz and Subik (1981) using lysing enzymes at 3 to 5 mg/mL. Spheroplasts were lysed and vacuoles isolated on a Ficoll step gradient according to Stevens et al. (1982). The vacuole fraction and the original spheroplast lysate were assayed for invertase and marker enzyme activities. α-Mannosidase was assayed according to Opheim (1978), NADPH cytochrome c reductase was assayed as described by Kubota et al. (1977), and α-glucosidase was assayed according to Halvorson and Elias (1958).

ACKNOWLEDGMENTS

REFERENCES

Akazawa, T., and Hara-Nishimura, I. (1985). Topographic aspects of biosynthesis, extracellular secretion, and intracellular storage of proteins in plant cells. Annu. Rev. Plant Physiol. 36, 441–472.

Ammerer, G. (1983). Expression of genes using the ADC1 promoter. Methods Enzymol. 101C, 192–201.

Bassüner, R., Manteuffel, R., Müntz, K., Püchel, M., Schmidt, P., and Weber, E. (1983). Analysis of in vivo and in vitro globulin formation during cotelydon development of field beans (Vicia faba L. var. minor). Biochem. Physiol. Pflanzen. 178, 665–684.

Bassüner, R., Wobus, U., and Rapoport, T.A. (1984). Signal recognition particle triggers the translocation of storage globulin polypeptides from field beans (Vicia faba L.) across mammalian endoplasmic reticulum membranes. FEBS Lett. 166, 314–320.

Bäumlein, H., Wobus, U., Pustell, J., and Kafatos, F.C. (1986). The legumin gene family: Structure of a B type gene of Vicia faba and a possible legumin gene specific regulatory element. Nucl. Acids Res. 14, 2707–2720.

Bäumlein, H., Müller, A., Schiemann, J., Helbing, D., Manteuffel, R., and Wobus, U. (1987). A legumin B gene of Vicia faba is expressed in developing seeds of transgenic tobacco. Biol. Zentralbl. 106, 569–575.

Bäumlein, H., Müller, A., Schiemann, J., Helbing, D., Manteuffel, R., and Wobus, U. (1988). Expression of a Vicia faba legumin B gene in transgenic tobacco plants: Gene dosage-dependent protein accumulation. Biochem. Physiol. Pflanzen. 183, 205–210.

Bednarek, S.Y., Wilkins, T.A., Dombrowski, J.E., and Raikhel, N.V. (1990). A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. Plant Cell 2, 1145–1155.

Botstein, D., Falco, S.C., Stewart, S.E., Brennan, M., Scherer, S., Stinchcomb, D.T., Struhl, K., and Davis, P.W. (1979). Sterile host yeast (SHY): A eucaryotic system of biological containment for recombinant DNA experiments. Gene 8, 17–24.

Burgess, T.L., and Kelly, R.B. (1987). Constitutive and regulated secretion of proteins. Annu. Rev. Cell Biol. 3, 243–293.

Burns, D.K., and Crowl, R.M. (1987). An immunological assay for chloramphenicol acetyltransferase. Anal. Biochem. 162, 399–404.

Carlson, M., and Botstein, D. (1982). Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. Cell 28, 145–154.

Carlson, M., Osmond, B.C., and Botstein, D. (1981). Mutants of yeast defective in sucrose utilization. Genetics 98, 25–40.

Chrispeels, M.J. (1991). Sorting of proteins in the secretory system. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42, in press.

Chrispeels, M.J., Higgins, T.J.V., and Spencer, D. (1982). Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. J. Cell Biol. 93, 306–313.

Craig, S., and Goodchild, D.J. (1982). Post-embedding immunolabelling. Some effects of tissue preparation on the antigenicity of plant proteins. Eur. J. Cell Biol. 28, 251–256.

Croy, R.R.D., Derbyshire, E., Krishna, T.G., and Boulter, D. (1979). Legumin of Pisum sativum and Vicia faba. New Phytol. 83, 29–35.

Dorel, C., Voelker, T.A., Herman, E.M., and Chrispeels, M.J. (1989). Transport of proteins to the plant vacuole is not by bulk flow through the secretory system, and requires positive sorting information. J. Cell Biol. 108, 327–337.

Gabriel, O., and Wang, S. (1969). Determination of enzymic activity in polyacrylamide gels. I. Enzymes catalyzing the conversion of nonreducing substrates to reducing products. Anal. Biochem. 27, 545–554.

Goldstein, A., and Lampen, J.O. (1975). β-D-Fructofuranoside fructohydrolase from yeast. Methods Enzymol. 42, 504–511.

Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. 2, 1044–1051.

Greenwood, J.S., and Chrispeels, M.J. (1985). Correct targeting of the bean storage protein phaseolin in seeds of transformed tobacco. Plant Physiol. 79, 65–71.

Halvorson, H.O., and Ellas, L. (1958). The purification and properties of an α-glucosidase of Saccharomyces italicus Y1225. Biochim. Biophys. Acta 30, 28–40.

Hara-Nishimura, I., and Hara-Nishimura, M. (1987). Proglobulin processing enzyme in vacuoles isolated from developing pumpkin cotyledons. Plant Physiol. 85, 440–445.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., and Fraley, R.T. (1985). A simple and general method for transferring genes into plants. Science 227, 1229–1231.

Horstmann, C. (1983). Specific subunit pairs of legumin from Vicia faba. Phytochemistry 22, 1861–1866.

Hurkman, W.J., and Beevers, L. (1980). What is pea legumin—is it glycosylated? Planta 150, 82–88.

Ito, H., Fukada, Y., Murata, K., and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. 153, 163–168.

Johnson, L.M., Bankaitis, V.A., and Emr, S.D. (1987). Distinct sequence determinants direct intracellular sorting and modification of a yeast vacuolar protease. Cell 48, 875–885.

Jones, R.L., and Robinson, D.G. (1989). Protein secretion in plants. New Phytol. 111, 567–597.

Kelly, R.B. (1985). Pathways of protein secretion in eukaryotes. Science 230, 25–32.

Klionsky, D.J., Banta, L.M., and Emr, S.D. (1988). Intracellular sorting and processing of a yeast vacuolar hydrolase: Proteinase A propeptide contains vacuolar targeting information. Mol. Cell. Biol. 8, 2105–2116.

Kubota, S., Yoshida, Y., Kumaoka, H., and Furumichi, A. (1977). Studies on the microsomal electron-transport system of anaerobically grown yeast. J. Biochem. 81, 197–206.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

Maraz, A., and Subik, J. (1981). Transmission and recombination of mitochondrial genes in Saccharomyces cerevisiae after protoplast fusion. Mol. Gen. Genet. 181, 131–133.

Müller, M., Marti, T., and Kriz, S. (1980). Improved structural preservation by freeze-substitution. In Proceedings of the 7th European Congress of the Electron Microscopy Foundation, Leiden, 1980, Vol. 2, P. Brederoo and W. DePriester, eds, pp. 720–721.

Opheim, D.J. (1978). α-D-Mannosidase of Saccharomyces cerevisiae: Modulation of activity. Biochim. Biophys. Acta 524, 121–130.

Perara, E., and Lingappa, V.R. (1988). Transport of proteins into and across the endoplasmic reticulum membrane. In Protein Transfer and Organelle Biogenesis, R.C. Das and P.W. Robbins, eds (New York: Academic Press), pp. 3–47.

Perlman, D., Halvorson, H.O., and Cannon, L.E. (1982). Presecretory and cytoplasmic invertase polypeptides encoded by distinct mRNAs derived from the same structural gene differ by a signal sequence. Proc. Natl. Acad. Sci. USA 79, 781–785.

Pfeffer, S.R., and Rothman, J.E. (1987). Biosynthetic protein transport by the endoplasmic reticulum and Golgi. Annu. Rev. Biochem. 56, 829–852.

Roth, J. (1983). The colloidal gold marker system for light and electron microscopic cytochemistry. In Techniques in Immunocytochemistry, Vol. 2, G.R. Bullock and P. Petrusz, eds (New York: Academic Press), pp. 217–284.

Roth, J., Bendayan, M., Carlemalm, E., Villinger, W., and Garavito, M. (1981). Enhancement of structural preservation and immunocytochemical staining in low temperature embedded pancreatic tissue. J. Histochem. Cytochem. 29, 663–671.

Rothman, J.E. (1989). Polypeptide chain binding proteins: Catalysts of protein folding and related processes in cells. Cell 59, 591–601.

Saalbach, G., Jung, R., Saalbach, I., and Müntz, K. (1988). Construction of storage protein genes with increased number of methionine codons and their use in transformation experiments. Biochem. Physiol. Pflanzen. 183, 211–218.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Schekman, R. (1985). Protein localization and membrane traffic in yeast. Annu. Rev. Cell Biol. 1, 115–143.

Schlesier, B., Bassüner, R., Van Hai, N., and Müntz, K. (1990). The cDNA derived primary structure of two distinct legumin A subunit precursors of field bean (Vicia faba) seeds. Nucl. Acids Res. 18, 7146.

Schwaiger, H., Hasilik, A., von Figura, K., Wiemken, A., and Tanner, W. (1982). Carbohydrate-free carboxypeptidase Y is transferred into the lysosome-like yeast vacuole. Biochem. Biophys. Res. Commun. 104, 950–956.

Sherman, F., Fink, G.R., and Hicks, J.B. (1986). Methods in Yeast Genetics: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Sonnewald, U., Sturm, A., Chrispeels, M.J., and Willmitzer, L. (1989). Targeting and glycosylation of patatin, the major potato tuber protein in leaves of transgenic tobacco. Planta 179, 171–180.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503–517.

Stevens, T.H., Esmon, B., and Schekman, R. (1982). Early stages in the yeast secretory pathway are required for transport of carboxypeptidase Y to the vacuole. Cell 30, 439–448.

Sturm, A., Voelker, T.A., Herman, E.M., and Chrispeels, M.J. (1988). Correct glycosylation, Golgi-processing, and targeting to protein bodies of the vacuolar protein phytohemagglutinin in transgenic tobacco. Planta 175, 170–183.

Tague, B.W., and Chrispeels, M.J. (1987). The plant vacuolar protein, phytohemagglutinin, is transported to the vacuoles of transgenic yeast. J. Cell Biol. 105, 1971–1979.

Tague, B.W., Dickinson, C.D., and Chrispeels, M.J. (1990). A short domain of the plant vacuolar protein phytohemagglutinin targets invertase to the yeast vacuole. Plant Cell 2, 533–546.

Tanaka, A., Ohnishi, N., and Fukui, S. (1967). Studies on the formation of vitamins and their function in hydrocarbon fermentation. J. Ferment. Technol. 45, 617–623.

Taussig, R., and Carlson, M. (1983). Nucleotide sequence of the yeast SUC2 gene for invertase. Nucl. Acids Res. 11, 1943–1954.

Thomas, P.S. (1983). Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. Methods Enzymol. 100, 255–266.

Tooze, J., Kern, H.F., Fuller, S.D., and Howell, K.E. (1989). Condensation-sorting events in the rough endoplasmic reticulum of exocrine pancreatic cells. J. Cell Biol. 109, 35–50.

Töpfer, R., Pröls, M., Schell, J., and Steinbiss, H.-H. (1987). A set of plant expression vectors for transcriptional and translational fusions. Nucl. Acids Res. 15, 5890.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. Procedure and some applications. Proc. Natl. Acad. Sci. USA 76, 4350–4354.

Valls, L.A., Hunter, C.P., Rothman, J.H., and Stevens, T.H. (1987). Protein sorting in yeast: The localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide. Cell 48, 887–897.

Valls, L.A., Winther, J.R., and Stevens, T.H. (1990). Yeast carboxypeptidase Y vacuolar targeting signal is defined by four propeptide amino acids. J. Cell Biol. 111, 361–368.

Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M., and Schell, J. (1983). Intergeneric transfer and exchange recombination of restriction fragments cloned in pBR322: A novel strategy for reversed genetics of Ti plasmids of Agrobacterium tumefaciens. EMBO J. 2, 411–417.

Vervliet, G., Holsters, M., Teuchy, H., Van Montagu, M., and Schell, J. (1975). Characterization of different plaque-forming and defective temperate phages in Agrobacterium strains. J. Gen. Virol. 26, 33–48.

Vieira, J., and Messing, J. (1987). Production of single-stranded plasmid DNA. Methods Enzymol. 153, 3–11.

Voelker, T.A., Herman, E.M., and Chrispeels, M.J. (1989). In vitro mutated phytohemagglutinin genes expressed in tobacco seeds: Role of glycans in protein targeting and stability. Plant Cell 1, 95–104.

von Figura, K., and Hasilik, A. (1986). Lysosomal enzymes and their receptors. Annu. Rev. Biochem. 55, 167–193.

von Schaewen, A., Stitt, M., Schmidt, R., Sonnewald, U., and Willmitzer, L. (1990). Expression of a yeast-derived invertase in the cell wall of tobacco and Arabidopsis plants leads to accumulation of carbohydrate and inhibition of photosynthesis and strongly influences growth and phenotype of transgenic tobacco plants. EMBO J. 9, 3033–3044.

Wickerham, L.J. (1946). A critical evaluation of the nitrogen assimilation tests commonly used in the classification of yeasts. J. Bacteriol. 52, 293–301.

Wieland, F.T., Gleason, M.L., Serafini, T.A., and Rothman, J.E. (1987). The rate of bulk flow from the endoplasmic reticulum to the cell surface. Cell 50, 289–300.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., van Montagu, M., and Schell, J. (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. EMBO J. 2, 2143–2150.

Zoller, M.J., and Smith, M. (1983). Oligonucleotide directed mutagenesis of DNA fragments cloned into M13 derived vectors. Methods Enzymol. 190, 468–500.

Zur Nieden, U., Manteuffel, R., Weber, E., and Neumann, D. (1984). Dictyosomes participate in the intracellular pathway of storage proteins in developing Vicia faba cotyledons. Eur. J. Cell Biol. 34, 9–17.

Transgenic tobacco plants expressing yeast-derived invertase in either the cytosol, vacuole or apoplast: a powerful tool for studying sucrose metabolism and sink/source interactions

Uwe Sonnewald[1][*], Monika Brauer[2], Antje von Schaewen[1][†], Mark Stitt[2] and Lothar Willmitzer[1]

[1]Institut für Genbiologische Forschung Berlin GmbH, Ihnestrasse 63, 1000 Berlin 33 and

[2]Lehrstuhl für Pflanzenphysiologie, Universität Bayreuth, Universitätsstrasse 30, 8580 Bayreuth, Germany

Summary

In higher plants sucrose plays a central role with respect to both short-term storage and distribution of photoassimilates formed in the leaf. Sucrose is synthesized in the cytosol, transiently stored in the vacuole and exported via the apoplast. In order to elucidate the role of the different compartments with respect to sucrose metabolism, a yeast-derived invertase was directed into the cytosol and vacuole of transgenic tobacco plants. This was in addition to the targeting of yeast-derived invertase into the apoplast described previously. Vacuolar targeting was achieved by fusing an N-terminal portion (146 amino acids long) of the vacuolar protein patatin to the coding region of the mature invertase protein.

Transgenic tobacco plants expressing the yeast-derived invertase in different subcellular compartments displayed dramatic phenotypic differences when compared to wild-type plants. All transgenic plants showed stunted growth accompanied by reduced root formation. Starch and soluble sugars accumulated in leaves indicating that the distribution of sucrose was impaired in all cases. Expression of cytosolic yeast invertase resulted in the accumulation of starch and soluble sugars in both very young (sink) and older (source) leaves. The leaves were curved, indicating a more rapid cell expansion or cell division at the upper side of the leaf. Light-green sectors with reduced photosynthetic activity were evenly distributed over the leaf surface. With the apoplastic and vacuolar invertase, the phenotypical changes induced only appear in older (source) leaves. The development of bleached and/or necrotic sectors was linked to the source state of a leaf. Bleaching followed the sink to source transition, starting at the rim of the leaf and moving to the base. The bleaching was paralleled by the inhibition of photosynthesis.

Introduction

Photosynthesis represents the major source for energy used to support biological processes in all living organisms. As primary products of photosynthesis, carbohydrates are formed in photosynthetically active cells and organisms. In higher plants, leaves and to a certain extent also other parts of the plant, e.g. stem tissue, represent the primary sites for photosynthesis. In contrast, other parts of the plant, e.g. roots, seeds or tubers, do not contribute significantly to the whole energy gain reached via photosynthesis but rather are largely dependent on carbon dioxide fixed in other photosynthetically active parts of the plant. Thus there is a net flow of energy from photosynthetically active tissues, representing the sources (defined as net exporters of fixed carbon), to photosynthetically inactive tissues of the plant, representing the sinks (defined as net importers of fixed carbon).

The primary products of carbon fixation are starch and sucrose. Whereas starch in leaves mainly serves as an intermediate deposit for the products of carbon fixation, sucrose, in addition to representing a storage form for the products of photosynthesis, plays a central role in the distribution of photoassimilates throughout the plant, especially the supply of photoassimilates to sinks.

Due to its dual role, sucrose is present in several different compartments. Sucrose is synthesized in the cytosol and is transiently stored in the vacuole. Finally, sucrose is probably exported from the leaf by transferring it into the apoplast followed by active uptake into the phloem (Turgeon, 1989). Although the involvement of the different compartments with respect to sucrose metabolism and/or distribution has been known for many years, little is known about the role of the different compartments in relation to sucrose metabolism.

Here we describe a new means of approaching this problem. This method takes advantage of our ability to create transgenic plants containing foreign genes, thus allowing the modulation of endogenous genes and/or ectopic expression of alien gene products. These plants can be disturbed at a particular step in the biosynthesis, and the storage and/or distribution of certain metabolites

95

can then be analyzed by various biochemical and physiological means and compared to near-isogenic wild-type plants. This makes it possible to follow the influence of this disturbance at the molecular level.

In order to introduce such disturbances into the metabolism of sucrose, we decided to express a foreign (yeast-derived) invertase in the cytosol and vacuole of transgenic plants in addition to its expression in the apoplast described previously (von Schaewen *et al.*, 1990). This invertase should give rise to cleavage of sucrose yielding glucose and fructose and thus should interfere with the normal biosynthesis, storage and distribution of sucrose. The yeast invertase was chosen for two reasons: first, this enzyme should not be inhibited by endogenous invertase inhibitors present in plants, and secondly, this enzyme is known to be active over a broad pH range (Goldstein and Lampen, 1976), which is important in relation to the different pH values of various subcellular compartments.

Whereas secretion of foreign proteins has already been described in transgenic plants (Dorel *et al.*, 1989; von Schaewen *et al.*, 1990), no data have as yet been reported on the targeting of chimeric non-plant-derived proteins into the vacuole of transgenic plants. Secreted or vacuolar proteins are synthesized as pre-proteins with a hydrophobic N-terminal signal peptide. In plants it has been shown that the signal peptide is required for secretion of chimeric proteins. Further information is needed for the targeting of foreign proteins into the vacuole (Dorel *et al.*, 1989). In the case of β-1,3 glucanases it has been speculated that vacuolar and secreted forms differ by C-terminal extensions being present in the cDNAs of the vacuolar but not the secreted proteins (Van den Bulcke *et al.*, 1989). Furthermore, removal of the C-terminal propeptide of the barley lectin gene led to the secretion of the truncated protein in transgenic tobacco plants, indirectly demonstrating the vacuolar targeting function of the C-terminal propeptide (Bednarek *et al.*, 1990). In yeast, however, it was shown that certain N-terminal regions of either plant-derived or yeast-derived vacuolar proteins are sufficient for achieving vacuolar targeting of chimeric proteins (Tague *et al.*, 1990; Valls *et al.*, 1987). Here we describe how the fusion of a large N-terminal-derived portion of the vacuolar protein patatin to the mature invertase protein from yeast results in vacuolar targeting of the invertase. To the best of our knowledge, this represents the first example for directing of a foreign protein into the vacuoles of a transgenic plant.

The transgenic plants expressing the yeast-derived invertase either in the cytosol, the vacuole or the apoplast show a clear phenotype which varies with the compartment in which the invertase is expressed and also depends on the developmental stage of the plant. Despite the fact that plants expressing the invertase in the different compartments show distinct differences, there are certain similarities on both the visual and the biochemical phenotype. These experiments therefore prove the central importance of the compartmentation of sucrose with respect to its biosynthesis, storage and distribution. Furthermore, these plants demonstrate the power of reversed genetics for analyzing and understanding a variety of physiological problems.

Results

Construction of chimeric genes directing yeast invertase into the cytosol or the vacuole

Eukaryotic cells develope specialized compartments to separate different anabolic and metabolic functions. To direct newly synthesized proteins to their proper destination, cellular receptor proteins and specific targeting signals of newly synthesized proteins interact. To manipulate biochemical pathways in any given organelle the isolation of suitable enzymes as well as targeting signals is required.

Proteins are synthesized in the cytosol and no further information is needed if they are to remain there. Cytosolic accumulation of yeast invertase was achieved via a fusion between the 5'-untranslated leader of the proteinase inhibitor II (PI-II) gene from potato and the yeast suc 2 coding region (subsequently called *Cy-INV*, Figure 1).

Secreted or vacuolar proteins are synthesized as pre-proteins with a hydrophobic N-terminal signal peptide. As described above, in the case of yeast it has been demonstrated that targeting signals needed to achieve vacuolar localization of foreign proteins are present in the N-terminal portion of either plant- or yeast-derived vacuolar proteins. We therefore decided to test whether or not this would also hold true for chimeric genes in plants. Therefore, a
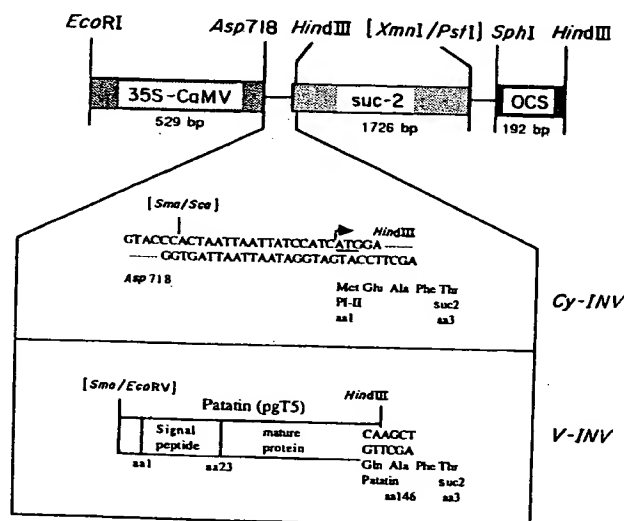

Figure 1. Structure of the cytosolic (*Cy-INV*) and the vacuolar (*V-INV*) invertase gene.

DNA fragment of the patatin genomic clone pgT5 (Rosahl et al., 1986), containing parts of the 5'-untranslated leader, the 23 amino acid signal peptide and also the 123 N-terminal amino acids from the mature patatin protein, was fused with the suc 2 gene from yeast (Figure 1) in order to achieve vacuolar targeting of the chimeric protein (subsequently called V-INV).

Chimeric yeast invertase is active in all three compartments

The chimeric yeast invertase genes were transformed into tobacco plants using Agrobacterium-mediated gene transfer. Transgenic tobacco plants arising after kanamycin selection were analyzed for the presence of intact non-rearranged genes by DNA gel blotting (data not shown). Only plants containing intact copies of the chimeric genes were used for further analysis. In each case, about 50 independent transformants were tested for invertase activity using the invertase gel assay system. Five plants of each of the individual invertase constructs displaying different amounts of the alien invertase activity were transferred to the greenhouse for further analysis.

In order to determine whether the different invertase constructs were active after targeting them to different compartments, total protein was isolated from leaves of transgenic tobacco plants and separated on SDS–PAGE. As shown in Figure 2, invertase activity could be detected under semi-native conditions in SDS–PAGE in all three compartments. It is interesting to note that the cytosolic invertase migrates with much higher mobility than the secreted and vacuolar forms. This difference cannot be explained only by the fact that the cytosolic form is not glycosylated whereas the two other forms are highly glycosylated (data not shown), but probably also results from the fact that only the glycosylated form produces oligomeric aggregates, whereas the cytosolic form is present as a monomer, as deduced from its migration (cf. Esmon et al., 1987). During leaf development an increase in invertase activity is found for the vacuolar and the secreted

invertase, whereas the activity of the cytosolic form decreases (Table 1). A further decrease in the activity of the cytosolic invertase occurs during plant ageing.

A fusion protein between patatin and yeast invertase is efficiently targeted to the vacuole of transgenic tobacco

The correct targeting of the patatin–invertase fusion protein into plant vacuoles was verified using three different experimental techniques.

First, transient expression experiments were carried out where the V-INV gene was introduced into Arabidopsis mesophyll protoplasts. After a cultivation period of 2–3 days, protoplasts were separated from the cultivation medium and both fractions were individually assayed for invertase activity (Figure 3). The patatin–invertase fusion protein was only detectable in protoplasts (Figure 3, lane 5) and was not secreted into the culture medium (Figure 3, lane 6). When constructs expressing the secreted chimeric yeast invertase were introduced into protoplasts as a control, invertase activity in both the protoplasts and culture medium increased. Invertase activity only increased in the protoplasts when constructs expressing the cytosolic invertase were introduced into protoplasts as controls. Untransformed control protoplasts did not show any invertase activity under the experimental conditions used (Figure 3, lanes 7 and 8).

The transient expression data alone showed that the V-INV-encoded protein was not secreted but stayed within the protoplasts. The V-INV-encoded protein was completely retained on a concanavalin A (ConA) Sepharose column, whereas the Cy-INV-encoded protein did not bind to ConA. This suggests a glycosylation of the V-INV-encoded protein, which therefore makes a cytosolic location unlikely, but a location within the endoplasmic reticulum–Golgi pathway has to be assumed.

Cell fractionation via differential centrifugation of Arabidopsis protoplasts after transient expression demonstrated that the patatin-invertase fusion protein was not precipitated with the endoplasmic reticulum–Golgi vesicles

Table 1. Effect of leaf age on invertase activity

Leaves	Sector	Invertase activity (μmol sucrose m^{-2} sec^{-1})			
		Wild-type	Cy-INV	Cw-INV	V-INV
Young[a]	Green	0.54 ± 0.04	25.79 ± 2.66	11.02 ± 0.44	3.55 ± 0.62
Middle[b]	Green	0.20 ± 0.01	5.09 ± 0.61	26.50 ± 8.47	4.90 ± 1.51
	Light green	—	6.67 ± 1.36	39.44 ± 13.26	7.00 ± 0.55
Old[c]	Green	0.16 ± 0.05	1.71 ± 0.16	42.99 ± 7.30	5.22 ± 1.72
	Light green	—	1.38 ± 0.12	56.67 ± 3.52	6.00 ± 0.56

[a] Leaf length less than 10 cm.
[b] Leaf length between 10 and 15 cm.
[c] Leaf length between 20 and 25 cm.
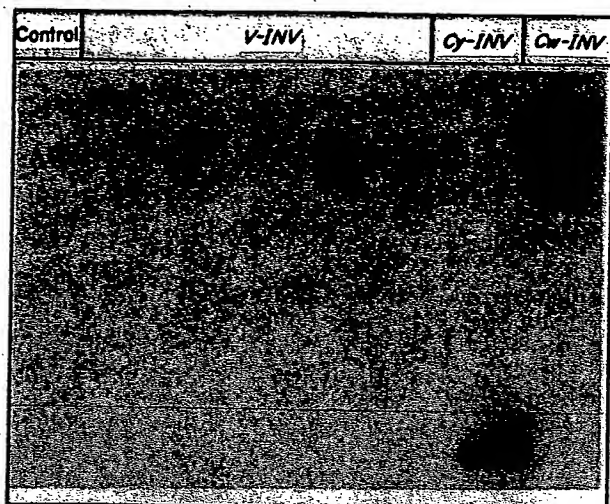
Figure 2. SDS–PAGE separation showing that chimeric yeast invertase is active in the cytosol, vacuole and apoplast of transgenic tobacco.
Protein extracts (50 µg total soluble protein each lane) were separated by SDS–PAGE (omitting the boiling step before electrophoretic separation, leaving active and intact invertase aggregates) and assayed for invertase activity (see Experimental Procedures). No invertase activity is detectable in nontransformed tobacco plants (control). In transformed tobacco plants a new invertase activity is detectable for the vacuolar (*V-INV*), the cytosolic (*Cy-INV*) and the cell wall (*Cw-INV*) transformants.

(Figure 4, lane P3) but was present in the soluble fraction (Figure 4, lane Sol).

As a final proof for the localization of the *V-INV*-encoded protein, vacuoles from transgenic tobacco mesophyll protoplast were isolated. After their isolation, the activity of the vacuolar marker enzyme α-mannosidase was determined for intact protoplasts and isolated vacuoles. The same amount of α-mannosidase activity for both protoplasts and vacuoles was loaded on SDS–PAGE and the invertase activity determined. As seen in Figure 5, invertase activity was detectable in isolated vacuoles to the same extent as in intact protoplasts. These experiments clearly demonstrate that the chimeric yeast invertase segregates with the vacuolar marker enzyme α-mannosidase, thus confirming its vacuolar localization.

Expression of chimeric yeast invertase in the cytosol, the vacuole and the apoplast leads to the accumulation of carbohydrates in leaves of transgenic tobacco

Accumulation of yeast invertase in different compartments of the cell is expected to interfere with the synthesis, transport and/or storage of sucrose. Consequently, changes in the carbohydrate pools of transgenic plants were investigated. In all cases starch accumulated in leaves and was not degraded during the dark period (Table 2), indicating that the leaf was not able to transport the fixed carbohydrate to other parts of the plant. Surprisingly, sucrose increased in all the transformants. This increase has been consistently
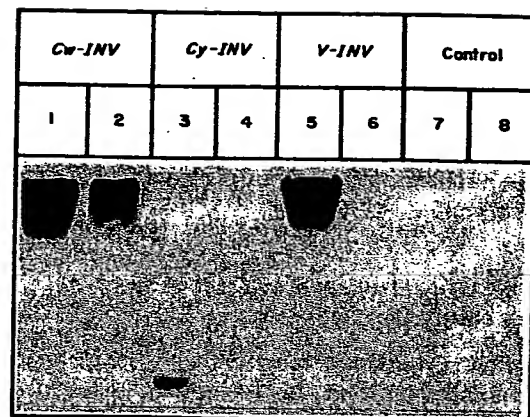
Figure 3. Transient expression of the different chimeric invertase genes in *Arabidopsis thaliana* protoplasts.
Mesophyll protoplasts isolated from axenically grown *Arabidopsis* plants were transformed with plasmid DNA of the different invertase constructs. After 72 h protoplasts were pelleted and the proteins present in the supernatant (lanes 2, 4, 6 and 8) and the pellet (lanes 1, 3, 5 and 7) were separated by SDS–PAGE and subsequently assayed for invertase activity. Lanes 1 and 2, cell wall invertase (it is important to note that in other experiments less invertase was detectable in the protoplasts, which might be explained by a different capability of the protoplasts to secrete proteins). Lanes 3 and 4, cytosolic invertase.
Lanes 5 and 6, vacuolar invertase.
Lanes 7 and 8, untransformed control protoplasts.
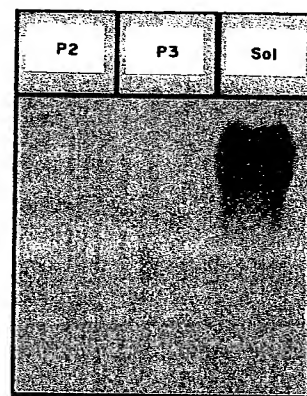
Figure 4. Cell fractionation of *Arabidopsis* protoplasts after transient expression of the *V-INV* gene.
Following transient expression cells were lysed and subjected to a series of centrifugation steps. After removal of cell debris (700 *g* for 5 min) the supernatant was centrifuged at 10 000 *g* for 10 min, yielding pellet P2 (containing large organelles). The resulting supernatant was subjected to a high-speed centrifugation (80 000 *g* for 2 h) to pellet membranes and vesicles (pellet P3). The membrane-free supernatant (Sol) contains both the cytoplasmic and the vacuolar cell fraction. All fractions were brought to the same volume and 50 µl of each fraction were subjected to SDS–PAGE. The gel was developed for invertase activity.

observed for the cell wall invertase in three different groups of tobacco plants (see for example, von Schaewen *et al.*, 1990; Stitt *et al.*, 1990), as well as in *Arabidopsis* transformants (von Schaewen *et al.*, 1990). In other experiments, we have seen that sucrose sometimes increases (see
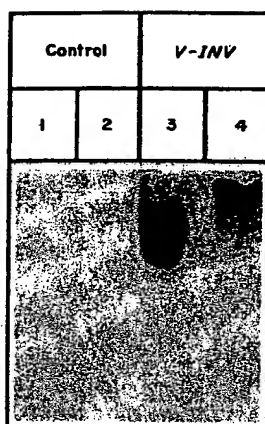
Figure 5. Detection of chimeric invertase in the vacuoles of transgenic tobacco.
Vacuoles were isolated from mesophyll protoplasts of transgenic tobacco plants carrying the *V-INV* gene and assayed for invertase activity after gel electrophoresis. The invertase activity in the vacuole was compared to the invertase activity present in intact protoplasts after loading the same amount of α-mannosidase activity (equivalent to 250 ng of 'jack bean' α-mannosidase; see Experimental Procedures) onto each lane.
Untransformed control: lane 1, protoplasts; lane 2, isolated vacuoles.
Transgenic tobacco expressing the *V-INV* gene: lane 3, protoplasts; lane 4, isolated vacuoles.

Table 2), but may remain unaltered or, in the case of the transformants expressing the vacuolar or cytosolic form, decrease slightly (M. Brauer, unpublished results). The experiments reported in Table 2 were carried out with plants growing in a greenhouse during the summer, whereas experiments in which sucrose did not accumulate were carried out with plants from growth chambers kept at lower irradiances.

Differences were found in the relative amounts of glucose and fructose (Table 2). The cleavage of sucrose via invertase action would lead to glucose and fructose in a ratio of 1:1, which is indeed found in the case of cell wall invertase. Cleavage of sucrose within the cytosol leads to a 5–10-fold larger accumulation of fructose than glucose. This excess of fructose (sometimes being as large as 40-fold) has consistently been observed in experiments on

three separate groups of plants. In the case of the vacuolar transformant, there was a 2–4-fold excess of glucose over fructose in these plants (Table 2) but this difference was not present in other experiments carried out with plants from a growth chamber (data not shown).

Accumulation of carbohydrates leads to drastic changes in the development and habit of transgenic tobacco plants

During tissue culture, transgenic tobacco plants expressing yeast invertase in the different compartments were mainly characterized by a reduced growth compared to untransformed control plants. This mild phenotype changed dramatically after transfer of the transgenic plants to soil and their subsequent growth in the greenhouse. Although the expression of the different invertase constructs resulted in distinct phenotypes they had some features in common. As a consequence of the invertase activity the plants showed a reduced height compared to control plants (Figure 6a). The reduced height was accompanied by reduced leaf expansion (data not shown) and root development (Figure 6e). The reduced growth of the vegetative organs led to late flowering and reduced seed setting (data not shown). Thus decreasing the amount of sucrose available for distribution prolonged the life of the transgenic plants for several months.

In addition to the phenotypic changes mentioned above, transformants harboring the cytosolic yeast invertase were characterized by curled leaves (Figures 6b and 7b). The curling started early in development and was seen on very young leaves. Evenly distributed light-green sectors also developed on the leaf surface and these bleached completely during leaf ageing.

Symptom development of plants carrying the cell wall and the vacuolar yeast invertase depends on the developmental stage of the leaf. Very young leaves showed no phenotypical changes. In the case of the apoplastic yeast invertase, bleached and/or necrotic sectors developed on mature source leaves following different patterns (described

Table 2. Carbohydrate turnover in invertase-expressing transgenic tobacco plants

| | Carbohydrate turnover (mmol hexose m^{-2}) | | | | | | | |
| | Dark[a] | | | | Light[b] | | | |
Genotype	Glucose	Fructose	Sucrose	Starch	Glucose	Fructose	Sucrose	Starch
Wild-type	0.18	0.12	0.21	0.44	0.22	0.16	0.77	9.19
Cy-INV	0.47	6.7	0.44	7.0	1.1	4.72	1.4	16.4
Cw-INV	9.8	8.1	4.5	8.3	10.9	9.8	11.7	10.3
V-INV	14.6	5.85	1.9	16.4	19.8	4.03	4.23	19.5

[a] Measured after 16 h darkness.
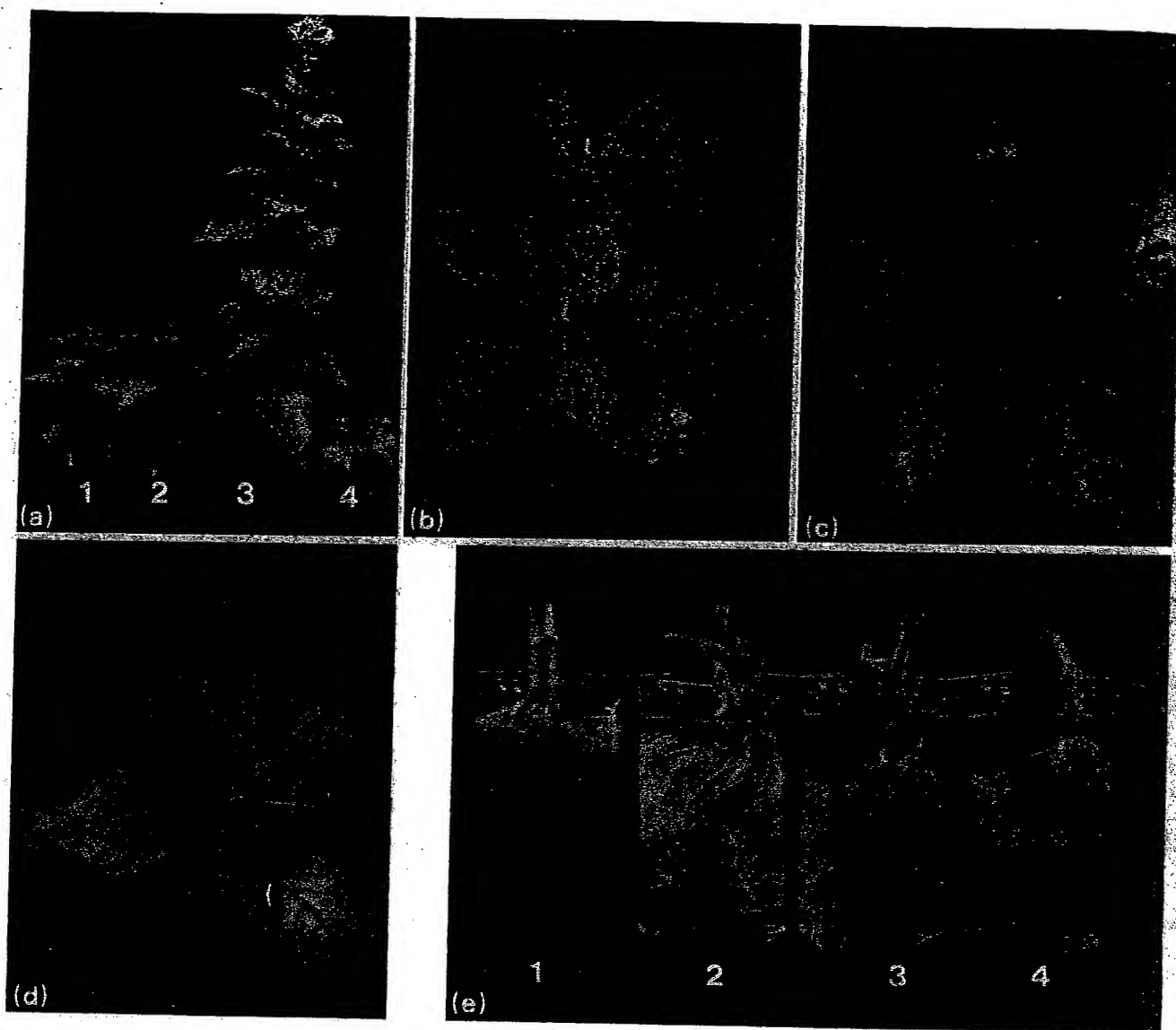[b] Measured after 8 h illumination.

Figure 6. Influence of chimeric yeast invertase on development and habit of transgenic tobacco.
(a) Stunted growth of transgenic tobacco plants expressing yeast invertase in the cytosol (*Cy-INV*; 1), the cell wall (*Cw-INV*; 2) and the vacuole (*V-INV*; 4) compared with the growth of a wild-type tobacco plant (3).
The development of symptoms in leaves of transgenic tobacco is distinct for each mutant, *Cy-INV* (b), *Cw-INV* (c) and *V-INV* (d).
(e) Impaired root formation of invertase-expressing tobaco plants: (1) wild-type, (2) *Cy-INV*; (3) *Cw-INV*; (4) *V-INV*.

in von Schaewen *et al.*, 1990). In transgenic plants expressing the vacuolar invertase, bleached sectors always followed the sink–source transition zone. Symptoms first became visible at the rim of the leaf and moved towards the base during leaf expansion (Figures 6c and 7a). It is important to note that, although the phenotypes of the cell wall and the vacuolar yeast invertase are quite similar, much less vacuolar invertase activity is needed to reduce the viability of transgenic tobacco plants.

The specific phenotype described for each of the different invertase constructs was consistently observed for all plants tested in the greenhouse. This excludes somaclonal variation as the cause for the changed habit. The phenotype was transmitted to the progeny and clearly linked to the appearance of the alien invertase activity. Furthermore, the severity of the phenotype correlated with the expression of the foreign invertase. Plants of the F_1 progeny of a *V-INV* transformant which showed only a mild phenotype in the parental generation are shown in Figure 8. A comparison of the severity of the phenotype with the invertase activity suggests that the invertase activity must be above a certain threshold value (≥ 4.5 μmol sucrose m^{-2} sec^{-1})
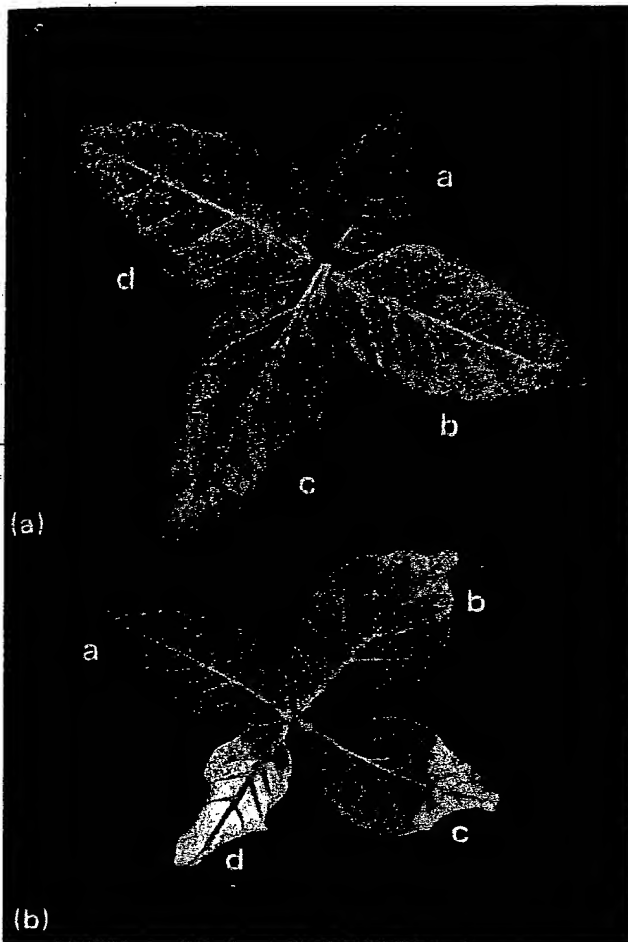
Figure 7. Development of symptoms is associated with the developmental stage of a leaf.
(a) Clockwise arrangement of leaves taken from a transgenic tobacco plant expressing cytosolic yeast invertase: a–d, young to old leaves.
(b) Clockwise arrangement of leaves taken from a vacuolar-invertase expressing transgenic tobacco plant: a–d, young to old leaves.

Figure 8. The development of symptoms follows the expression of the *V-INV* gene in the F_1 progeny of transgenic tobacco plants.
The progeny of transformant *V-INV-65* was tested for soluble, neutral invertase activity and the development of the phenotype was followed in the greenhouse.
The activities were: plant 1, 0.1 μmol sucrose m^{-2} sec^{-1}; plant 2, 5.5 μmol sucrose m^{-2} sec^{-1}; plant 3, 4.5 μmol sucrose m^{-2} sec^{-1}; plant 4, 2.8 μmol sucrose m^{-2} sec^{-1}.

(cf. also Figure 6d and Table 1) to influence development of the plant.

Photosynthesis is reduced in all transgenic plants expressing chimeric invertase proteins

As described above, expression of yeast invertase in transgenic tobacco plants leads to bleached areas, due to the loss of chlorophyll (data not shown). The percentage of bleached regions increased during leaf maturation (Table 3). In order to determine the rate of photosynthesis in relation to symptom development, leaves were separated into green, light-green and pale areas. As shown in Table 4, the rate of photosynthesis follows the development of symptoms. In young leaves (< 10 cm length) there was no

visual phenotype in plants expressing the cell wall invertase or vacuolar invertase and photosynthesis was even slightly higher than in the wild-type, whereas plants expressing the cytosolic invertase already showed an inhibition of photosynthesis even in those sectors of the smaller leaves which were visually unaffected (Table 4). As the leaves mature the appearance of a visual phenotype in the cell wall- or vacuolar-invertase expressing plants was accompanied by the inhibition of photosynthesis. This inhibition was restricted to the visually affected areas of the larger leaves (Table 4). The green areas had a similar or even higher rate of photosynthesis than the wild-type (see also von Schaewen *et al.*, 1990; Stitt *et al.*, 1990). The inhibition of photosynthesis in cytosolic invertase plants (relative to wild-type leaves of a comparable size) did not increase in the older leaves.

Table 3. Development of symptoms in invertase-expressing transgenic tobacco plants[a]

Genotype	Leaf length (cm)	Percentage of total leaf area for each phenotype		
		Green	Light-green	Pale
Cy-INV	20–25	—	—	—
	15–20	22.5	77.5	—
	10–15	35.0	65.0	—
	< 10	45.8	54.2	—
Cw-INV	20–25	32.6	57.7	9.7
	15–20	82.8	14.3	2.9
	10–15	89.3	10.7	—
	< 10	98.8	1.2	—
V-Inv	20–25	—	—	—
	15–20	47.9	40.9	11.2
	10–15	87.4	12.6	—
	< 10	100.0	—	—

[a] Development of symptoms was determined using a 5 × 5 mm screen and counting the visually affected areas.

The overall rate of photosynthesis of a leaf will depend on (i) the percentage of the area which has developed a phenotype (Table 3), (ii) the inhibition of photosynthesis (Table 4) and (iii) the size of the leaf (Table 4). As only part of the leaf is affected and the green areas retain high or enhanced rates of photosynthesis, the average rate per unit area remains relatively high (Table 5). However, the total yield of photosynthesis will be decreased due to the fact that the leaves of invertase transformants are smaller.

Table 5. Overall rate of photosynthesis[a]

Leaf length (cm)	Rate of photosynthesis (percentage of wild-type)		
	Cy-INV	Cw-INV	V-INV
20–25	—	83.9	79.1
15–20	63.6	112.0	93.4
10–15	49.4	99.9	85.1
< 10	57.3	106.0	115.0

[a] The rate of photosynthesis was determined for the different sectors of invertase-bearing leaves. The overall rate of photosynthesis per unit area was calculated after counting the portion of each individual sector.

Discussion

Targeting of chimeric yeast invertase into the vacuole

In order to apply the tools of reversed genetics to physiological problems, efficient ways of targeting foreign proteins into different subcellular compartments are needed because many biochemical processes are highly compartmentalized. Protein synthesis takes place in the cytosol and from here the proteins have to be targeted to their proper destination. The transport of proteins into mitochondria or chloroplasts is mediated by the N-terminal transit peptide (Chua and Schmidt, 1979). The first step in the secretion of proteins is the uptake of the precursor protein into the lumen of the endoplasmic reticulum (Blobel, 1980). In the absence of

Table 4. Rate of photosynthesis in different transgenic tobacco plants[a]

Genotype	Leaf length (cm)	Steady-state rate of photosynthesis for each phenotype (μmol O_2 m^{-2} sec^{-1})		
		Green	Light-green	Pale
Wild-type	25–30	25.25 ± 0.83	—	—
	20–25	24.79 ± 1.95	—	—
	15–20	24.94 ± 3.25	—	—
	10–15	27.47 ± 3.63	—	—
	< 10	17.78 ± 3.70	—	—
Cy-INV	15–20	30.75 ± 3.85	11.53 ± 1.27	—
	10–15	19.20 ± 6.56	10.56 ± 2.27	—
	< 10	14.52 ± 1.35	6.50 ± 2.10	—
Cw-INV	20–25	30.18 ± 2.88	18.34 ± 0.85	4.21 ± 2.42
	15–20	30.03 ± 1.32	17.57 ± 1.15	—
	10–15	27.50 ± 2.87	27.00 ± 0.89	—
	< 10	18.85 ± 1.65	—	—
V-INV	15–20	24.57 ± 3.15	18.04 ± 2.91	3.98 ± 1.89
	10–15	23.41 ± 1.64	23.34 ± 1.92	—
	< 10	20.45 ± 1.4	—	—

[a] The rate of photosynthesis was determined at 20°C, saturating CO_2 and 600 μmol photons m^{-2} sec^{-1}.

retention (Munro and Pelham, 1987) or vacuolar targeting signals, the proteins will be secreted along the default pathway (Dorel *et al.*, 1989). Targeting to the vacuole is not well understood. Studies on the N-glycosylation of vacuolar proteins (Sonnewald *et al.*, 1990; Voelker *et al.*, 1989) demonstrated that the targeting signal must be part of the polypeptide chain of vacuolar proteins. In the case of the barley lectin, the C-terminal propeptide was shown to be necessary for proper targeting of this lectin to vacuoles of transgenic tobacco plants (Bednarek *et al.*, 1990). To date there have been no reports in the literature of the successful targeting of chimeric proteins to the vacuole of higher plants.

In order to obtain vacuolar targeting of yeast invertase, we have fused the first 123 N-terminal amino acids of the mature patatin protein to the suc 2 coding region. Our results are the first example of the targeting of a non-vacuolar protein to vacuoles of transgenic tobacco. Furthermore, the fusion protein provides enzymatic activity and can be used in physiological studies designed to investigate the role of sucrose in the different cellular compartments.

Alteration in invertase activity as the leaf develops

The constructs were all under the control of the 35S-CaMV promoter so that changes in their rate of accumulation are presumably due to the stability of the invertase proteins in the different compartments. Cell wall protein (von Schaewen *et al.*, 1990) accumulated during leaf development and reached the highest levels of activity compared to the other invertase forms, although the amount of detectable RNA was lowest (data not shown). Cytosolic invertase activity was highest in young leaves and decreased as the leaf matured. This might indicate that different and/or more proteases are expressed during leaf development. The activity of the vacuolar invertase was found to be rather low compared with the other forms, even though the amount of RNA was highest (data not shown), which might indicate a lower stability of the fusion protein in the vacuole.

The amount of invertase with which the plant can cope depends on its subcellular location

As expected, the plant is very sensitive to invertase in the cytosol, and less sensitive to invertase in the cell wall. The cytosolic invertase will directly interfere with sucrose synthesis, whereas cell wall invertase will only be effective when it is in the right place to interfere with the passage of sucrose to the phloem. It is surprising that tobacco is so sensitive to an invertase in the vacuole. This indicates that the flux of carbon between the cytosol and the vacuole can be quite high. Rapid turnover of the vacuolar pool of

sucrose has been reported previously for pea and barley (Borland and Farrar, 1988; Farrar and Farrar, 1986).

The time-course of the visual phenotype depends on the site at which invertase is expressed

The plants with cytosolic invertase already showed their full phenotype in young leaves. This shows that sucrose synthesis is obligatory in young leaves, even if they do not export sucrose. The sucrose is presumably used as a temporary storage product and can be regulated far better than free hexoses.

In plants with cell wall or vacuolar invertase the phenotype developed later, probably during the sink–source transition. This is easily explained in plants with cell wall invertase, because from the time of this transition onwards the cell wall invertase will gain ready access to sucrose moving through the apoplast to the phloem (von Schaewen *et al.*, 1990). Since, in the case of the cytosolic invertase transformant, the phenotype had already developed in young leaves (see above), we conclude that the phenotype is caused by a perturbation of the sucrose metabolism and accumulation in the leaf, rather than other developmental changes associated with the sink–source transition *per se*.

The delayed development of the phenotype in plants with vacuolar invertase is unexpected. The activity of invertase did not increase greatly during leaf expansion. One possible explanation would be that the permeability of the tonoplast for sucrose increases during the later stages of leaf development. This could represent a mechanism to allow leaves to stop storing sucrose in the vacuole (for osmotic or short-term storage purposes) and start exporting it. Another, although much more speculative, possibility is that sucrose has to pass through the vacuole in order to become exported.

Changes of carbohydrate depend on the subcellular location of the invertase

In all cases, invertase resulted in increased levels of free hexoses. For the cell wall invertase, glucose and fructose were present in approximately equal amounts (see also von Schaewen *et al.*, 1990; Stitt *et al.*, 1990), as expected from the hydrolysis of sucrose. The vacuolar transformant showed a 2–4-fold excess of glucose over fructose in these experiments, although this difference has not always been observed. One possible explanation is that the vacuolar transformant accumulated large amounts of starch; we have observed in other experiments that wild-type tobacco often contains an excess of glucose when the leaves have a high starch content (W.P. Quick, personal communication). However, more experiments are required

because several other explanations are also possible, including preferential transport of fructose out of the vacuole, or accumulation of fructans in the vacuolar transformants.

In contrast, fructose accumulated to a 5–10-fold excess over glucose in cytosolic invertase plants. This has been observed consistently in other experiments. The accumulation of fructose could be explained either by the absence of sufficient fructokinase activity, or by inhibition of fructokinase in these plants. Preliminary experiments (M. Brauer, unpublished data) indicate that these plants contain fructokinase but that the enzyme is very susceptible to product inhibition by fructose, as has also been reported for fructokinase from soybean nodules (Copeland and Morell, 1985) and maize kernels (Doehlert, 1989). Tobacco leaves also contain a broad-specificity hexokinase, but this enzyme has a fivefold lower affinity for fructose than for glucose (M. Brauer, unpublished data). These results could explain the observed imbalance between glucose and fructose in the cytosolic invertase transformant.

This explanation, if correct, implies that most of the hexoses in the cell wall or vacuolar transformant are located in the apoplast or vacuole, i.e. after increasing the activity of invertase dramatically in these compartments, the transport of hexoses back into the cytosol may have become limiting. Accumulation of hexoses outside the cytosol could also result in osmotic problems. In this context, it is interesting that vacuolar and cell wall transformants share the phenotype feature that the leaves remain green and do not accumulate sugars in the vicinity of the major vascular bundles. The supply of water in and around the vascular bundles will be large and water flow in the transpiration stream would also (in case of the cell wall invertase) tend to move sugars in the apoplast away from the vascular bundle and concentrate them in the intervening lamellar sectors (von Schaewen et al., 1990; Stitt et al., 1990). This provides an explanation for the mosaic phenotype observed in these plants.

One rather surprising feature of our results is that the sucrose concentration in the leaves of the invertase transformants often increases. This has been observed consistently for the cell wall transformants, and occasionally for the cytosolic and vacuolar transformants. The variation in the case of the latter two transformants may depend on the growth conditions, with sucrose accumulating under high light conditions. More studies are required, but our results indicate that disruption of sucrose metabolism by invertase may sometimes lead to direct inhibition of transport, in addition to effects caused by hydrolysis of sucrose *en route* to the phloem. We have already shown that the inhibition of photosynthesis is caused by a general reduction of photosynthetic enzymes (Stitt et al., 1990) and it might be speculated that the expression of proteins needed for sucrose transport is decreased in these transformants.

Sink development is affected by invertase

In addition to inhibiting photosynthesis, all of the transformants showed inhibited leaf and root development. This can be explained because sucrose export from the source leaves will be decreased in all these transformants. However, it is striking that the largest inhibition of leaf expansion rate was observed in the cytosolic transformant, and that these leaves were also visually curled and thick. It might be speculated that sink leaves on the cytosolic transformant cannot store sucrose properly, and that normal growth is disturbed because of osmotic effects exerted by the high concentrations of fructose in the young leaves.

The curling of the leaves might be explained by a higher hexose content on the upper side of the leaf (compared to the lower side), due to higher photosynthetic activity of these cells, which would lead to increased water influx, resulting in a more rapid cell expansion or even cell division. This shows that disruption of sink–leaf sucrose metabolism affects leaf expansion and development.

In conclusion, we have transformed tobacco plants with yeast invertase and targeted the protein to the cell wall, the cytosol, or the vacuole. The transformants show a clear development-specific phenotype which varies depending on the compartment in which the invertase is expressed. There are similarities, but also striking differences, in the biochemical phenotype in the three transformants. These experiments therefore directly demonstrate the central importance of compartmentation and protein targeting in plant development and metabolism. Furthermore, these plants provide a novel and powerful tool to investigate a variety of physiological problems, including the 'sink' regulation of photosynthesis by accumulating carbohydrate (von Schaewen et al., 1990; Stitt et al., 1990), the co-ordination of the metabolism and transport of sucrose, osmoregulation, and the regulation of leaf development.

Experimental procedures

Plants, bacterial strains and media

Nicotiana tabacum L. 'Samsun NN' was obtained through 'Vereinigte Saatzuchten' (Ebstorf, Germany). *Arabidopsis thaliana* L. 'Columbia C24' was kindly provided by J.P. Hernalsteens (Vrije Universiteit, Brussels, Belgium). Plants in tissue culture were grown under a 16 h light/8 h dark regime on Murashige and Skoog medium (Murashige and Skoog, 1962) containing 2% sucrose (2MS). Plants used for biochemical analysis were grown in the greenhouse as described in von Schaewen et al. (1990). *Escherichia coli* strain DH5a (Bethesda Research Laboratories, Gaithersburg, USA) was cultivated using standard techniques (Maniatis et al., 1982). *Agrobacterium tumefaciens* strain C58C1 containing pGV2260 (Debleare et al., 1985) was cultivated in YEB medium (Vervliet et al., 1975).

Reagents

DNA restriction and modification enzymes were obtained from Boehringer Mannheim (Ingelheim, Germany) and New England Biolabs (Danvers, USA). Synthetic oligonucleotides were synthesized on an Applied Biosystems (Foster City, USA) DNA Synthesizer (380A). Reagents for SDS–PAGE were purchased from BioRad (St Louis, USA). Chemicals were obtained through Sigma Chemical Co. (St Louis, USA) or Merck (Darmstadt, Germany).

Plasmid construction

In order to obtain cytosolic invertase activity, the 400 bp *Asp*718/ *Hin*dIII fragment of the plasmid PI-58-INV (von Schaewen *et al.*, 1990) was exchanged with a synthetic oligonucleotide (Figure 1). This exchange resulted in the removal of the PI-II signal sequence (Keil *et al.*, 1986), while the untranslated leader was unchanged. Vacuolar invertase was constructed using a DNA fragment (nucleotides 707–1895) from the patatin genomic clone pgT5 (Rosahl *et al.*, 1986) containing parts of the untranslated leader, the signal sequence, and also 597 nucleotides coding for the mature protein which allows vacuolar targeting of fusion proteins. This DNA fragment was fused via a *Hin*dIII-linker to the suc 2 gene containing nucleotides +64 to +1765 (Taussig and Carlson, 1983). The chimeric gene obtained was introduced as an *Eco*RV/ *Sph*I fragment into a plant expression cassette containing the 35S promoter of cauliflower mosaic virus (CaMV) (nucleotides 6909–7437; Franck *et al.*, 1980) and the polyadenylation signal of the octopine synthase gene (OCS). The constructs were cloned into the binary vector pBin19 (Bevan, 1984) as *Eco*RV/*Hin*dIII fragments and directly transformed into *Agrobacterium* strain C58C1:pGV2260.

Transient expression in Arabidopsis *protoplasts*

Isolation, transformation and cultivation of *Arabidopsis* protoplasts was essentially as described by Damm *et al.* (1989). The processing of the protoplasts and the culture medium for further experiments followed the protocol of von Schaewen *et al.* (1990).

Cell fractionation of Arabidopsis *protoplasts after transient expression*

Transient expression experiments were carried out as described previously (von Schaewen *et al.*, 1990). After a 2–3 day cultivation period, 4×10^6 cells were pelleted for 5 min at 80 g in a microfuge tube. The supernatant was removed with a drawn-out Pasteur pipet. The cells were resuspended in 100 μl of cold sucrose buffer, which is isotonic for intracellular organelles (12% sucrose [w/w], 1 mM DTT (DL-dithiothreitol), 1 mM EDTA, 50 mM Tris–HCl pH 7.8, 1 μg f.c. protease inhibitor cocktail, i.e. Leupeptin, Antipain and Pepstatin), and lysed by vortexing in the presence of glass beads (five times for 30 sec at room temperature, beads to volume ratio 1:4). The extraction mixture was then subjected to a series of centrifugation steps to remove cellular debris (700 g for 5 min, yielding P1) and large organelles, e.g. chloroplasts and mitochondria (10 000 g for 10 min, yielding P2). The supernatant of P2 was centrifuged (Beckman 50Ti, Beckman Instruments, Palo Alto, USA) for 2 h at 80 000 g, at 4°C to pellet membranes and vesicles (P3). The membrane-free supernatant of P3 was referred to as the soluble mixture of both the cytoplasmic and the vacuolar cell fractions (Sol). Fractions P2, P3 and Sol were brought to the same volume (100 μl) with sucrose buffer and 50 μl aliquots were subjected to SDS–PAGE.

Invertase activity

Detection of yeast invertase after SDS–PAGE and measurement of alkaline invertase activity was as described by von Schaewen *et al.* (1990). Proteins were applied to the gel without prior boiling in the presence of SDS thus leaving protein aggregates intact and active.

Tobacco transformation

Transformation of tobacco plants was carried out using the *Agrobacterium tumefaciens* leaf disc technique as described by Rosahl *et al.* (1987).

Isolation of plant vacuoles from protoplasts

Vacuoles were released from leaf mesophyll protoplasts by a combination of osmotic and thermal shock, using a modification of the protocol described by Boller and Kende (1979). Pelleted protoplasts ($3–5 \times 10^6$) were chilled on ice shortly before resuspension in 6 ml of prewarmed (40°C) lysis medium (0.2 M mannitol, 10% Ficoll Type 400, 20 mM EDTA, 2 mM DTT, 5 mM Hepes, pH 8.0) and transferred into a 15 ml Corex tube at room temperature. Within 5–10 min, release of vacuoles from the lysed protoplasts (followed under the microscope) was complete. The mixture was briefly chilled on ice and overlayered with 3 ml of cold 4% Ficoll solution (1 volume lysis medium + 1.5 volumes vacuole buffer) followed by 1 ml of cold vacuole buffer (0.45 M mannitol, 10 mM Hepes pH 7.5, 1 mM L-cysteine, 1 μg ml^{-1} each of Leupeptin, Antipain and Pepstatin). The gradient was centrifuged for 30 min at 5000 g in a 'swing-out' rotor at 10°C. Under these conditions cellular debris pellets, unlysed protoplasts collect at the 4% Ficoll interphase and vacuoles float to the top of the gradient. The vacuoles were collected with a Pasteur pipet from the miniscus of the 0% Ficoll step and then frozen at −70°C. The yield of vacuoles was usually 20% of the initially lysed protoplasts, as monitored microscopically by counting and enzymatically by α-mannosidase assays (Van der Wilden *et al.*, 1980). To compare invertase activity in protoplasts and the vacuolar fraction, equal α-mannosidase activities (equivalent to 250 ng α-mannosidase) were subjected to SDS–PAGE and stained for invertase activity as described previously (von Schaewen *et al.*, 1990). As a standard, α-mannosidase from 'jack bean' (Sigma) was used.

Measurement of photosynthesis

Eight leaf discs (each 0.5 cm²) were removed from tobacco plants and placed in a leaf disc O_2 electrode (Hansatech, Kings Lynn, UK). Illumination was 600 μmol photons m^{-2} sec^{-1} and saturating CO_2 was supplied from 400 μl 2 M $KHCO_3$/K_2CO_3 buffer pH 9.3 (Quick *et al.*, 1989) at 20°C. After 10 min illumination the samples were frozen in liquid N_2 under continuing illumination.

Determination of soluble sugars and starch

Leaf discs were taken at the indicated times and extracted with 80% ethanol (10 mM Hepes–KOH pH 7.4) at 80°C for 15 min. The supernatant was used for the determination of glucose, fructose and sucrose (Stitt *et al.*, 1989). The remaining leaf disc material was extracted a second time, washed in water and dried. Determination of starch was carried out as described by Stitt *et al.* (1978).

References

Bednarek, S.Y., Wilkins, T.A., Dombrowski, J.E. and Raikhel, N.V. (1990) A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. *Plant Cell*, **2**, 1145–1155.

Bevan, M. (1984) Binary *Agrobacterium* vectors for plant trans-formation. *Nucl. Acids Res.* **12**, 8711–8721.

Blobel, G. (1980) Intracellular protein topogenesis. *Proc. Natl Acad. Sci. USA*, **77**, 1496–1500.

Boller, T. and Kende, H. (1979) Hydrolytic enzymes in the central vacuole of plant cells. *Plant Physiol.* **63**, 1123–1132.

Borland, A.M. and Farrar, J.F. (1988) Compartmentation and fluxes of carbon in leaf blades and leaf sheaths of *Poa annua* L. and *Poa jemtlandica*. *Plant Cell Environ.* **11**, 535–543.

Chua, N.-H. and Schmidt, G.W. (1979) Transport of proteins into the mitochondria and chloroplasts. *J. Cell. Biol.* **81**, 461–483.

Copeland, L. and Morell, M.K. (1985) Hexokinases from the plant cytosolic fraction of soybean nodules. *Plant Physiol.* **79**, 114–117.

Damm, B., Schmidt, R. and Willmitzer L. (1989) Efficient trans-formation of *Arabidopsis thaliana* using direct gene transfer to protoplasts. *Mol. Gen. Genet.* **217**, 6–12.

Deblaere, R., Bytebier, B., De Greve, H., Debroeck, F., Schell, J., van Montagu, M. and Leemans, J. (1985) Efficient octopine Ti plasmid-derived vectors of *Agrobacterium* mediated gene transfer to plants. *Nucl. Acids Res.* **13**, 4777–4788.

Doehlert, D.C. (1989) Separation and characterization of four hexose kinases from developing maize kernels. *Plant Physiol.* **89**, 1042–1048.

Dorel, C., Voelker, T., Herman, E.M. and Chrispeels, M.J. (1989) Transport of proteins to the plant vacuole is not by bulk flow through the secretory system and requires positive sorting information. *J. Cell. Biol.* **108**, 327–337.

Esmon, P.C., Esmon, B.E., Schauer, J.E., Taylor, A. and Schekman, R. (1987) Structure, assembly and secretion of octameric invertase. *J. Biol. Chem.* **262**, 4387–4394.

Farrar, S.C. and Farrar, J.F. (1986) Compartmentation and fluxes of sucrose in intact leaf blades of barley. *New Phytol.* **103**, 645–657.

Franck, A., Guilley, H., Jonard, G., Richards, K. and Hirth, L. (1980) Nucleotide sequence of Cauliflower Mosaic Virus DNA. *Cell*, **21**, 285–294.

Goldstein, A. and Lampen, J.O. (1975) β-D-Fructofuranoside fructohydrolase from yeast. *Methods Enzymol.* **42**, 504–511.

Keil, M., Sánchez-Serrano, J., Schell, J. and Willmitzer, L. (1986) Primary structure of a Proteinase Inhibitor II gene from potato (*Solanum tuberosum*). *Nucl. Acids Res.* **14**, 5641–5650.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Munro, S. and Pelham, H.R.B. (1987) A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, **48**, 899–907.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497.

Quick, W.P., Siegel, G., Neuhaus, H.E., Feil, R. and Stitt, M. (1989) Short-term water stress leads to a stimulation of sucrose synthesis by activating sucrose-phosphate synthase. *Planta*, **177**, 536–546.

Rosahl, S., Schmidt, R., Schell, J. and Willmitzer, L. (1986) Isolation and characterization of a gene from *Solanum tuberosum* encoding patatin, the major storage protein of potato tubers. *Mol. Gen. Genet.* **203**, 214–220.

Rosahl, S., Schell, J. and Willmitzer, L. (1987) Expression of a tuber-specific storage protein in transgenic tobacco plants: demonstration of an esterase activity. *EMBO J.* **6**, 1155–1159.

von Schaewen, A., Stitt, M., Schmidt, R., Sonnewald, U. and Willmitzer, L. (1990) Expression of a yeast-derived invertase in the cell wall of tobacco and *Arabidopsis* plants leads to accumu-lation of carbohydrate and inhibition of photosynthesis and strongly influences growth and phenotype of transgenic tobacco plants. *EMBO J.* **9**, 3033–3044.

Sonnewald, U., von Schaewen, A. and Willmitzer, L. (1990) Expression of mutant patatin protein in transgenic tobacco plants: role of glycans and intracellular location. *Plant Cell*, **2**, 345–355.

Stitt, M., Wirtz, W. and Heldt, H.W. (1978) Pathway of starch breakdown in photosynthetic tissue of *Pisum sativum*. *Biochim. Biophys. Acta*, **544**, 200–214.

Stitt, M., Lilley, R.Mc.C., Gerhardt, R. and Heldt, H.W. (1989) Metabolite levels in specific cells and subcellular compartments of plant leaves. *Methods Enzymol.* **174**, 518–552.

Stitt, M., von Schaewen, A. and Willmitzer, L. (1990) 'Sink' regulation of photosynthetic metabolism in transgenic tobacco plants expressing yeast invertase in their cell wall involves a decrease of Calvin-cycle enzymes and an increase of glycolytic enzymes. *Planta*, **183**, 40–50.

Tague, B.W., Dickinson, C.D. and Chrispeels, M.J. (1990) A short domain of the plant vacuolar protein phytohemagglutinin targets invertase to the yeast vacuole. *Plant Cell*, **2**, 533–546.

Taussig, R. and Carlson, M. (1983) Nucleotide sequence of the yeast suc 2 gene for invertase. *Nucl. Acids Res.* **11**, 1943–1954.

Turgeon, R. (1989) The sink-source transition in leaves. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 119–138.

Valls, L.A., Hunter, C.P., Rothman, J.H. and Stevens, T.H. (1987) Protein sorting in yeast: the localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide. *Cell*, **48**, 887–897.

Van den Bulcke, M., Bauer, G., Castresana, C., Van Montagu, M. and Vanderkerckhove, J. (1989) Characterization of vacuolar and extracellular b(1,3)-glucanases of tobacco: evidence for a strictly compartmentalized system. *Proc. Natl Acad. Sci. USA*, **86**, 2673–2677.

Van der Wilden, W., Hermann, E.M. and Chrispeels, M.J. (1980) Protein bodies of mung bean cotyledons as autophagic organelles. *Proc. Natl Acad. Sci. USA*, **77**, 428–432.

Vervliet, G., Holsters, M., Teuchy, H., Van Montagu, M. and Schell, J. (1975) Characterization of different plaque-forming and defective temperate phages in *Agrobacterium* strains. *J. Gen. Virol.* **26**, 33–48.

Voelker, T., Herman, E.M. and Chrispeels, M.J. (1989) In vitro mutated phytohemagglutinin genes expressed in tobacco seeds: role of glycans in protein targeting and stability. *Plant Cell*, **1**, 95–104.

Expression of Rice Lectin Is Governed by Two Temporally and Spatially Regulated mRNAs in Developing Embryos

Thea A. Wilkins and Natasha V. Raikhel[1]

Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824-1312

Two cDNA clones encoding rice lectin have been isolated and characterized to investigate the expression of rice lectin at the molecular and cellular levels. The two cDNA clones code for an identical 23-kilodalton protein which is processed to the mature polypeptide of 18 kilodaltons by co-translational cleavage of a 2.6-kilodalton signal sequence and selective removal of a 2.7-kilodalton COOH-terminal peptide which contains a potential N-linked glycosylation site. In addition, the mature 18-kilodalton lectin is post-translationally cleaved between residues 94 and 95 to yield polypeptides of 10 kilodaltons and 8 kilodaltons, corresponding to the NH_2- and COOH-terminal portions of the mature subunit, respectively. RNA gel blot analysis established that rice lectin is encoded by two mRNA transcripts (0.9 kilobase and 1.1 kilobase). On DNA gel blots, the rice lectin cDNAs hybridize specifically to a single restriction fragment. In situ hybridization showed localization of the 1.1-kilobase rice lectin mRNA in root caps and specific cell layers of the radicle, coleorhiza, scutellum, and coleoptile. RNA gel blot analysis demonstrated that both the 0.9-kilobase and 1.1-kilobase mRNAs are present in developing rice embryos. The two lectin mRNAs are differentially expressed temporally such that the 1.1-kilobase lectin mRNA accumulates to levels twofold higher than the 0.9-kilobase mRNA.

INTRODUCTION

Plant lectins are a class of proteins that bind and cross-link specific carbohydrates. Because of their unique carbohydrate-binding properties, lectins are widely used as tools in medical cell biology (Lis and Sharon, 1986). Historically, plant lectin research has focused on the isolation and characterization of new lectin species to broaden the spectrum of specific carbohydrate-binding moieties. Although the function of lectins in plants remains obscure, dissecting the regulation of expression of lectin genes at the molecular level should facilitate elucidation of the protein function in vivo.

Many of the Gramineae synthesize N-acetylglucosamine (GlcNAc)-binding lectins with similar immunological properties (Peumans and Stinissen, 1983). These lectins accumulate in a cell-type specific manner in various organs of developing embryos and young seedlings. Rice lectin, initially purified and characterized by Tsuda (1979) from rice bran, is a dimeric protein composed of two glycine- and cysteine-rich 18-kD subunits that lack covalently bound sugar residues. In the cultivated rice species *Oryza sativa*, the majority of the 18-kD subunits undergo a pro-

teolytic cleavage event which yields two subunits of 8 kD and 10 kD (Stinissen, Peumans, and Chrispeels, 1984). This lectin is synthesized as a 23-kD monomeric precursor on the rough endoplasmic reticulum (RER) and is subsequently assembled into dimers within the lumen of the RER (Stinissen, Peumans, and Chrispeels, 1984). Assembled dimers are only transiently associated with the RER before being transported to and deposited in vacuoles/protein bodies (Stinissen, Peumans, and Chrispeels, 1984). Rice lectin accumulates in specific cell layers of the scutellum, coleorhiza, radicle, root cap, and throughout cell layers of the coleoptile of embryos (Mishkind, Palevitz, and Raikhel, 1983).

We are interested in the molecular mechanisms regulating cell-specific expression of the Gramineae lectins. Two cDNA clones encoding rice lectin have been isolated and used to examine the expression of rice lectin in developing embryos. In this paper, we present evidence that both cDNA clones represent two mRNA transcripts. Each lectin mRNA transcript exhibits a distinct pattern of temporal expression in developing embryos. Moreover, the cell-type specific expression of rice lectin mRNAs is developmentally and spatially regulated.

[1] To whom correspondence should be addressed.

lectin is a glutamine, which is presumably modified by cyclization to pyrrolidone carboxylic acid (Chapot et al., 1986), a residue resistant to Edman degradation. Initiating with glutamine Q + 1, the cDNAs encode a protein comprising 199 amino acids with calculated M_r 20,172. However, amino acid sequence analysis of the COOH-terminal amino acids indicates that mature rice lectin terminates at the glycine residue G173 (arrow in Figures 1A and 1B). Thus, determination of terminal amino acid residues revealed that the mature polypeptide of rice lectin comprises 173 amino acids with M_r 17,512 (arrows in Figures 1A and 1B and a solid box in Figure 1B). The amino acid composition indicates that the mature subunit of rice lectin is a glycine- and cysteine-rich polypeptide. Cysteine (23%) and glycine (19.7%) together account for almost 43% of the mature polypeptide amino acid composition.

In addition to the signal sequence and the mature rice lectin subunit, the cDNAs encode proteins with an additional 26 amino acids (M_r 2678) extending beyond the COOH terminus of mature rice lectin (boxed residues in Figure 1A, stippled box in Figure 1B). This COOH-terminal extension is a relatively hydrophobic domain and contains a potential N-linked glycosylation site at asparagine residue N179 (asterisk in Figures 1A and 1B). Rice lectin is therefore synthesized as a preproprotein that requires the proteolytic removal of the signal sequence and post-translational processing of a COOH-terminal domain to yield the mature polypeptide. In vacuoles, the mature 18-kD subunit polypeptide undergoes additional post-translational processing to yield two smaller polypeptides of approximately 10 kD and 8 kD (Stinissen, Peumans, and Chrispeels, 1984). To resolve the relationship between these polypeptides and the protein encoded by the cDNAs, both polypeptides were purified and subjected to NH₂-terminal and COOH-terminal amino acid sequence analyses. Results from these analyses indicate that the mature subunit of rice lectin is proteolytically cleaved between amino acids residues N94 and G95 as deduced from the cDNA clones (open arrowhead, Figure 1A). The resultant 10-kD and 8-kD polypeptides correspond to the NH₂- and COOH-terminal portions of the mature 18-kD protein, respectively.

A comparison of amino acids from rice lectin and isolectin B of wheat germ agglutinin (WGA-B) is presented in Figure 2. Rice lectin exhibits 73% identity with WGA-B (boxed amino acids in Figure 2) within the coding region of the mature subunits spanning from glutamine Q + 1 to glycine G171 in WGA-B or glycine G173 in rice lectin. The overall homology between the two lectins increases to 79.5% when conserved amino acid changes (asterisks in Figure 2) are included in the comparison. Both rice lectin and WGA-B require the post-translational processing of COOH-terminal domains to produce the mature 18-kD subunit. Alignment of the 26-amino acid COOH-terminal domain from the proprotein of rice lectin and the 15-amino acid COOH-terminal domain from pro-WGA-B for maximal
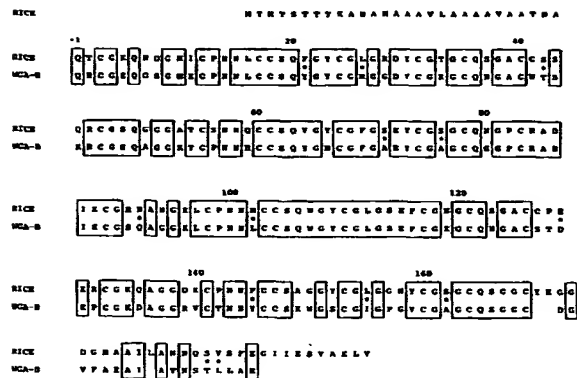
Figure 2. Comparison of Amino Acid Sequences between Rice Lectin and Isolectin B of Wheat Germ Agglutinin (WGA-B).

The complete deduced amino acid sequence of rice lectin was aligned for maximal homology to the available amino acid sequence of WGA-B (Raikhel and Wilkins, 1987). Identical amino acids are depicted by boxed residues, whereas conserved amino acid changes between the two lectins are denoted by asterisks.

homology shows a 46.7% overall amino acid conservation, indicating that this region is less conserved than the coding region of the mature protein.

Rice Lectin Is Encoded by Two Different mRNAs

To explore the relationship between the two cDNA clones encoding rice lectin, an RNA gel blot containing total RNA from developing rice embryos [10 days to 20 days postanthesis (DPA)] was probed with ³²P-labeled insert from cRL852 or cRL1035. Two mRNA species of approximately 1.1 kb and 0.9 kb were identified (Figure 3A). Therefore, the two cDNA clones correspond to the two mRNAs and did not arise from cloning artifacts. To discriminate expression due solely to the 1.1-kb mRNA species, a clone-specific probe (cRL165) encompassing 165 bp of the 3'-untranslated region unique to cRL1035 was constructed (see Figure 1B). The specificity of cRL165 as a clone-specific probe for the cDNA cRL1035 and the 1.1-kb lectin mRNA transcript was confirmed by DNA gel and RNA gel blot analyses, respectively (data not shown).

To determine the number of rice lectin genes responsible for the two mRNAs, DNA gel blots containing restricted genomic DNA were hybridized with ³²P-labeled inserts from cDNA clones cRL852, cRL1035, or cRL165. Figure 3B shows a representative DNA gel blot depicting single restriction fragments of 11 kb and 15 kb detected in EcoRI- and HindIII-digested genomic DNA (cv. IR36), respectively,
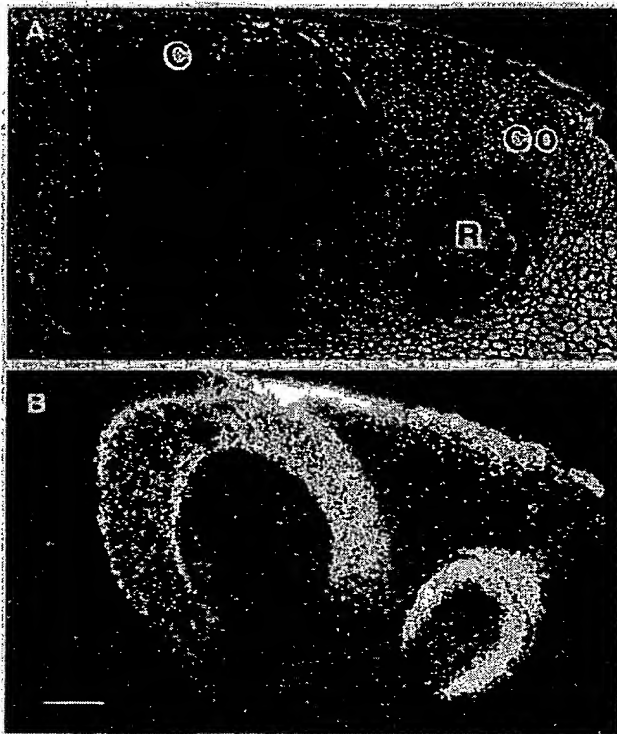
Figure 4. Localization of Rice Lectin mRNAs in a Developing Embryo.

In situ hybridization of embryo sections harvested at 20 DPA and hybridized with ^{35}S-labeled antisense transcript generated from the rice lectin cDNA clone cRL852 as described in Methods.
(A) Bright-field photomicrograph detailing the cellular organization of a developing rice embryo. C, R, Co, refer to the coleoptile, radicle, and coleorhiza, respectively.
(B) Dark-field photomicrograph depicting the tissue-specific distribution of rice lectin mRNA in a developing rice embryo. The white grains represent formation of RNA/RNA hybrids. Bar = 50 μm.

DISCUSSION

Two cDNA clones encoding the embryo-specific lectin of rice were isolated and used to characterize the expression of rice lectin in developing embryos at the molecular and cellular levels. The results from these studies have resolved the relationship between the different molecular forms of rice lectin. Previous studies demonstrated that the 23-kD precursor of rice lectin synthesized on the RER is post-translationally processed to an 18-kD polypeptide (Stinissen et al., 1984). However, the nature of these processing events, which result in the conversion of the 23-kD protein to the 18-kD form, remained unidentified. Characterization of two cDNAs, which encode identical 23-kD preproteins, indicates that processing of both NH[2]- and COOH-terminal sequences is required to generate the 18-kD rice lectin subunit. As predicted for proteins destined for entry into the endomembrane system of
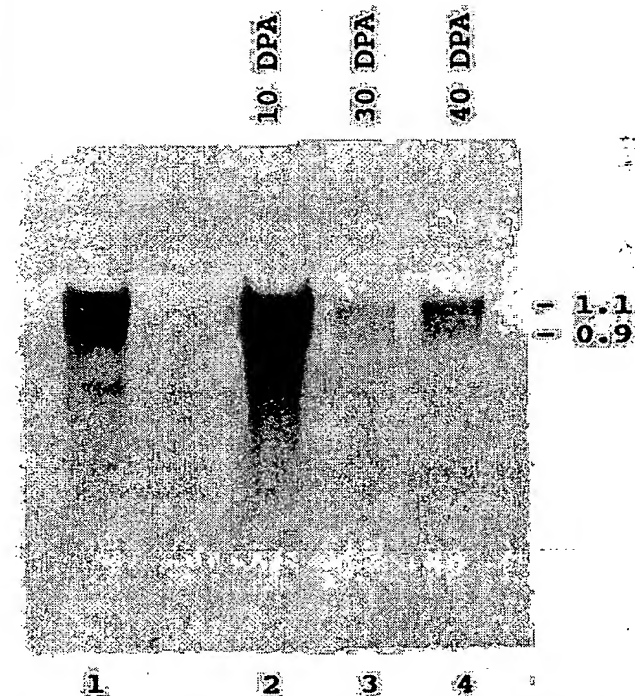
Figure 5. Differential Expression of Rice Lectin mRNA during Embryo Development.

RNA gel blot of total RNA from developing embryos was hybridized with ^{32}P-labeled insert from cDNA clone cRL852. Lane 1 contains 12.5 μg of total RNA from embryos harvested at 10 DPA. Lanes 2, 3, and 4 contain 25 μg of total RNA from embryos collected at 10, 30, and 40 DPA, respectively. Sizes (in kilobases) of each lectin mRNA are indicated to the right.

uration at around 20 DPA, relative levels of lectin mRNA present in total RNA decreased to low levels (Figure 5, lanes 3 and 4). The dramatic decline in lectin mRNA levels between 10 and 30 DPA embryos is accompanied by a temporal change in the relative levels of the individual mRNA species. In embryos at 10 DPA, the 0.9-kb and 1.1-kb lectin mRNA transcripts are present at similar levels. By 30 DPA, however, the 1.1-kb lectin mRNA accumulates to levels approximately twofold higher than the 0.9-kb mRNA. Thus, both mRNAs are expressed and differentially regulated in the radicles and coleoptiles of developing embryos.

METHODS

Plant Material

Developing rice (*Oryza sativa* cv. Lemont) embryos were collected from spikes harvested at 5 DPA, 10 DPA, 20 DPA, 30 DPA, and 40 DPA from plants maintained under greenhouse conditions. Embryos used for in situ hybridization experiments were processed immediately, while the bulk of collected embryos (10, 30, 40 DPA) were quick frozen in liquid nitrogen and stored at −80°C for RNA isolation.

Young seedlings of the rice cultivars Nato or IR36 were germinated and grown in Baccto professional potting mix in a growth chamber with a 12-hr light period at 27°C and a 12-hr dark period at 21°C, with 70% humidity. Shoots of 10-day-old seedlings were collected and frozen in liquid nitrogen for isolation of total DNA.

Screening of a λgt10 cDNA Library for Rice Lectin

A λgt10 cDNA library constructed from poly(A)⁺ RNA isolated from spikes of rice *O. sativa* cv. Nato was provided by Susan Wessler and Ron Okagaki (University of Georgia, Athens, GA). Approximately 160,000 recombinant phage were grown on *Escherichia coli* C600hfl at a density of 40,000 per 150-mm Petri plate and replicated onto nitrocellulose filters as described in Maniatis, Fritsch, and Sambrook (1982). The nitrocellulose filters were hybridized with a ^{32}P-random primer-labeled cDNA insert (Feinberg and Vogelstein, 1983) from clone WGA-B (clone pNVR1 described in Raikhel and Wilkins, 1987) for 18 hr in 6 × SSC, 5 × Denhardt's solution, 0.2% SDS, and sonicated salmon sperm DNA at 5 μg/ml. Post-hybridization washes included three 15-min washes at room temperature and two 15-min washes at 60°C in 3 × SSC, 0.1% SDS. Positive phage were plaque-purified to homogeneity (Maniatis, Fritsch, and Sambrook, 1982) under high stringency screening conditions using a ^{32}P-labeled insert from WGA-B (Raikhel and Wilkins, 1987).

DNA Nucleotide Sequence Analysis

Inserts, designated cRL852 and cRL1035, were purified from selected phage by electrophoresis in low-melting-point agarose (Struhl, 1985) and cloned into pUC119 (Vieira and Messing, 1987) in both orientations for subsequent DNA sequence determination. A sequential series of overlapping deletions from both strands of the cDNA was generated by T4 DNA polymerase (Dale and Arrow, 1987) from full-length, single-stranded DNA templates (Vieira and Messing, 1987). Single-stranded deletion templates were sequenced by the dideoxynucleotide chain termination method (Sanger, Nicklen, and Coulsen, 1977) using ^{35}S-dATP and 7-deaza-dGTP instead of dGTP (Mizusawa, Nishimura, and Seela, 1986). Computer alignment of overlapping deletions, and amino acid and sequence analysis were performed using Microgenie software (Beckman).

A fortuitous deletion encompassing the terminal 165 bp of 3′-untranslated region unique to cRL1035 was retrieved for use as a clone-specific probe. This partial cDNA clone was maintained in pUC119 and given the designation cRL165.

Amino Acid Sequence Determinations of NH$_2$- and COOH-Terminal Amino Acid Residues of Rice Lectin

Rice lectin was purified from 10 g of mature rice embryos (cv. IR36) via affinity chromatography on immobilized *N*-acetylglucosamine (Selectin 1, Pierce) according to the procedure detailed in Mansfield, Peumans, and Raikhel, (1988). To enhance resolution of the rice lectin during SDS-PAGE on a 15% polyacrylamide gel (Laemmli, 1970), the purified protein was S-carboxyamidated at 37°C for 30 min in the presence of 240 mM iodoacetamide (Raikhel, Mishkind, and Palevitz, 1984) prior to electrophoresis. Individual subunits (8 kD and 10 kD) of rice lectin were visualized by staining the gel briefly (10 min) in Coomassie blue, followed by destaining in 30% methanol, 7.5% acetic acid. The 8-kD and 10-kD polypeptides were excised from the gel, electroeluted in Laemmli (1970) buffer, and lyophilized. SDS was removed from protein by the organic extraction method of Konigsberg and Henderson (1983). Removal of salts from the protein was accomplished by dialysis against 15% acetic acid at 4°C in the dark for 2 days prior to lyophilization.

Approximately 200 pmol of gel-purified rice lectin was applied to a Model 477 Sequenator equipped with a 120 on-line PTH-amino acid analyzer (Applied Biosystems, Inc.) for determination of NH$_2$-terminal amino acid residues. The terminal amino acids of the COOH terminus were determined by carboxypeptidase Y digestion of 500 pmol of rice lectin via the procedure of Hayashi (1977). The identification and quantitation of free amino acids in digestion mixtures were accomplished by HPLC analysis using precolumn derivatization with *o*-phthaldialdehyde. Amino acid sequence determinations were performed at the Protein Chemistry Facility, University of California, Irvine.

RNA Gel Blot Analysis

Total RNA was isolated from 50 mg to 150 mg of developing rice embryos via the hot phenol method of Finkelstein and Crouch (1986) with the addition of 1% 2-mercaptoethanol to the homogenization buffer. RNA gel blots were prepared from 25 μg of RNA for each developmental stage of embryos and hybridized with random-primer-labeled cRL852 insert under stringent conditions (Raikhel, Bednarek, and Wilkins, 1988). Blots were exposed to Kodak XAR-5 film with intensifying screens at −80°C for 10 hr to 15 hr. Autoradiograms were scanned with a Gilford densitometer.

Gene Reconstruction Analysis

Total DNA was isolated from 10-day-old rice (cv. IR36 or Nato) seedlings according to Shure, Wessler, and Fedoroff (1983) and restricted to completion with EcoRI, HindIII, KpnI, SmaI, or XbaI. Two micrograms of digested DNA (3.3 × 10⁶ genome equivalents) and 0.5-copy, 1.0-copy, and 3.0-copy equivalents of the cRL1035 cDNA clone were fractionated by agarose gel electrophoresis and transferred to nitrocellulose (Maniatis, Fritsch, and Sambrook, 1982). Gene copy reconstructions were based upon a rice genome size of 5.47 × 10⁵ kb per haploid genome (Francis, Kidd, and Bennett, 1985). Hybridization and post-hybridization washes of the reconstruction blot were performed as described for RNA gel blots with the exception that random-primer radiolabeled insert from cRL1035 was used as a probe.

identification and isolation of the waxy locus in maize. Cell 35, 225–233.

Stinissen, H.M., Peumans, W.J., and Chrispeels, M.J. (1984). Subcellular site of lectin synthesis in developing rice embryos. EMBO J. 3, 1979–1985.

Struhl, K. (1985). A rapid method for creating recombinant DNA molecules. BioTechniques 3, 452–453.

Tsuda, M. (1979). Purification and characterization of a lectin from rice bran. J. Biochem. 86, 1451–1461.

Vieira, J., and Messing, J. (1987). Production of single-stranded plasmid DNA. Methods Enzymol. 153, 3–11.

von Heijne, G. (1983). Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. 133, 17–21.

Close similarity among streptavidin-like, biotin-binding proteins from *Streptomyces*

Edward A. Bayer *, Tikva Kulik, Rivka Adar, Meir Wilchek

Department of Biophysics, The Weizmann Institute of Science, Rehovot 76100, Israel

Abstract

Two strains of *Streptomyces venezuelae* were found to produce high-affinity, biotin-binding proteins, termed streptavidin v1 and v2, respectively. Both proteins were isolated to purity, and their corresponding genes were cloned and sequenced. Compared to streptavidin from *S. avidinii*, streptavidin v1 had only a single amino acid substitution and streptavidin v2 showed 9 such differences. The substitutions were remarkably conservative, none of which affected the amino acid residues known to be important to the biotin-binding properties or to the structure of the tetrameric protein. The results also indicate that the biosynthesis of such biotin-binding proteins is not simply a curious anomaly in a single species of *Streptomyces*. It is suggested that the classification of *S. avidinii* as a unique species should be reconsidered. The occurrence of these proteins appears to be linked to the production of an unusual synergistic antibiotic complex.

Keywords: Streptavidin; Sequence comparison; Biotin-binding protein; (*Streptomyces*)

1. Introduction

Egg-white avidin and bacterial streptavidin are two very similar biotin-binding proteins with very similar properties. In our early studies on the structure–function relationship of these two proteins, we performed chemical modification experiments [1–3] which were later supplemented by X-ray analysis [4]. Such studies indicated that there may be little room for drastic changes in their structure, since chemical modifications were usually associated with a loss of binding. This notion became even clearer when we sequenced an antibiotin antibody and found combining-site motifs similar to those of avidin and streptavidin [5].

However, chemical modification usually introduces bulky moieties into a protein, and the final conclusions emanating from such studies are often obscured. In order to overcome this problem, modification of amino acids which do not increase the size of a target amino acid are required. This can be accomplished by two different approaches: (i) by substituting specified amino acids by site-directed mutagenesis, and (ii) by comparing the structures of other biotin-binding proteins in nature.

We have decided to apply both approaches in our future studies. We begin with the second approach, under the premise that if significant changes in the binding site are detected, this would suggest which amino acids are suitable candidates for site-directed mutagenesis. If we find to the contrary, this would suggest that nature is indeed very conservative and any change in the binding site would probably lead to a reduction in the affinity for biotin.

The original discovery of streptavidin indicated that this protein forms part of a peculiar synergistic antibiotic complex in a novel species of *Streptomyces*, named *S. avidinii* [6,7]. More recently [8], a similar type of antibiotic activity was described in two strains of a different species (*S. venezuelae* strains Tü 2460 and Tü 2605). We therefore examined a variety of *Streptomyces* strains and found that only the two *S. venezuelae* strains which produce the antibiotic also produce biotin-binding proteins which are similar to streptavidin from *S. avidinii*. The respective proteins were termed streptavidin v1 and v2. Determination of their sequences revealed a very close correlation to the original streptavidin, indicating a strict conservation of the residues for biotin binding and protein structure.

---

* Corresponding author. E-mail: bfbayer@weizmann.weizmann.ac.il. Fax: +972 8 468256.

2. Materials and methods

2.1. Bacteria

Streptomyces venezuelae strains Tü 2460 and Tü 2605 were generous gifts of Professors H. Zähner and G. Jung of Tübingen University, Germany. Type strains of *S. avidinii* ATCC 27419, *S. cattleya* NRRL 8057, *S. clavuligerus* ATCC 27064 (DSM 738), *S. griseus* ATCC 10137, *S. jumonijinensis* NRRL 5641, *S. lavendulae* ATCC 8664 and *S. venezuelae* ATCC 10712 were obtained from the respective culture collections.

S. lividans 66, *S. lactamdurans* MA2908, *S. coelicolor* A317, and *Streptomyces* strain aleph 50 were supplied by Y. Aharonowitz of Tel Aviv University, Israel.

2.2. Materials

Restriction and modifying enzymes were purchased from Boehringer-Mannheim Biochemica (Mannheim, Germany) or from Promega (Madison, WI, USA). The DNA markers comprised the 1 kb ladder from Life Technologies (Gaithersburg MD, USA). Sequenase V2.0 was a product of United States Biochemical (Cleveland, OH, USA). *Escherichia coli* XL1-Blue was from Stratagene Cloning Systems (LaJolla, CA, USA). Radiochemicals and nylon filters (Hybond-N, 0.45 μm) were obtained from Amersham International (Buckinghamshire, UK). Common laboratory reagents, biochemicals and chemicals were obtained from Sigma (St. Louis, MO, USA) or from E. Merck (Darmstadt, Germany).

2.3. Hybridization

For genomic blotting, chromosomal DNA from different strains of *Streptomyces* was prepared as described previously [9]. The samples were digested for 4 h using the desired enzyme (e.g., *Bam*HI, *Pst*I, *Sma*I, and *Kpn*I).

The DNA digests were loaded on a 1% agarose gel equilibrated with TAE buffer (40 mM Tris-acetate and 1 mM EDTA, pH 8.0). Following electrophoresis (16 h, 7 mA), the DNA in the gels were denatured by 0.5 M NaOH in 1.5 M NaCl, neutralized using 1 M Tris (pH 7.4) in 1.5 M NaCl, and transferred to nylon filters [10]. The samples were crosslinked to the membrane by ultraviolet irradiation at 254 nm using a Stratagene crosslinker at an intensity of 120 mJ cm^{-2}.

Two oligonucleotide probes were designed. Probe 2492 (5'-ATG CAT ATG CGC AAG ATC GTC GTT GCA GC-3') corresponded to positions 50–72 of the streptavidin gene [11] with an *Nde*I restriction site attached at the N-terminus, and probe 2493 (5'-CTT AAG CTT CTA CTG AAC GGC GTC G-3' corresponded to positions 583–601 with a *Hind*III site at the C-terminus. The oligonucleotide probes were labeled at the 5' end using polynucleotide kinase and [γ-^{32}P]ATP [12].

Blots were prehybridized at 68°C for 4 h with a mixture of 5 × Denhardt's, 5 × SSC, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA. Hybridization was accomplished using the same treatment with 16 h incubation at 42°C. The blots were washed twice for 30 min at 25°C with 2 × SSC plus 0.1% SDS and similarly with 0.1% SSC plus 0.1% SDS.

2.4. Cloning of streptavidin genes

DNA samples from *S. avidinii* or from *S. venezuelae* Tü 2460 were digested with *Bam*HI, and the fragments were separated on 1% low-melting agarose gels. Fragments 2–2.5 kb in length were separated and ligated into *Bam*HI-digested plasmids (pGEM and pUC18, respectively). DNA from *S. venezuelae* Tü 2605 was digested with *Pst*I, and fragments of ~ 5000 kb were excised from the gels and ligated into pUC18 plasmid. *Escherichia coli* XL1-Blue was used as a host. Probes 2492 and 2493 were used for screening. Prehybridization and hybridization temperatures were 68°C and 42°C, respectively.

2.5. Sequencing

Nucleotide sequences were determined by the dideoxynucleotide chain-termination method [13] using Sequenase (United States Biochemical) according to the protocol supplied by the manufacturer. To complete the sequences, various synthetic oligonucleotide sequencing primers were used to cover both strands.

2.6. Detection of biotin-binding activity

The presence of biotin-binding proteins in the culture medium of the various strains of *Streptomyces* was determined colorimetrically on microtiter plates as described previously [14].

2.7. Preparation of streptavidin and analogs

Cells of *S. avidinii*, *S. venezuelae* Tü 2460 or 2605 were grown on malt medium, and the biotin-binding proteins in the culture broth were purified using an iminobiotin-Sepharose resin [15]. The purified proteins were dialyzed against distilled water and lyophilized. The biotin-binding activity of a weighed sample of each protein was determined by the HABA method [16].

Table 1
Isolation of biotin-binding proteins from *Streptomyces* strains

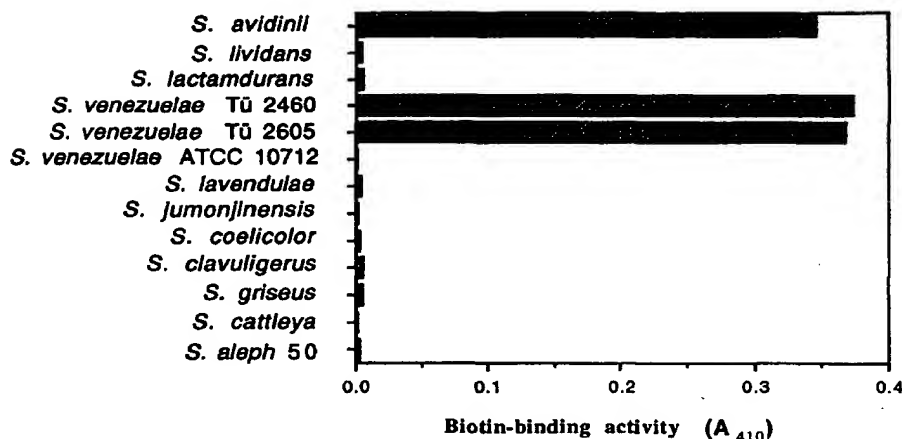| Strain	Protein	Yield (mg/l)
S. avidinii	streptavidin	12.5
S. venezuelae Tü 2460	streptavidin v1	24.0
S. venezuelae Tü 2605	streptavidin v2	22.0

Fig. 1. Secretion of biotin-binding activity in various species of *Streptomyces*. The designated strains were grown on malt medium and the cell-free growth medium was assayed for biotin-binding activity by an ELISA-like colorimetric assay system as described in the text.

2.8. Miscellaneous methods

SDS-PAGE of the streptavidins was performed on boiled samples using 15% gels according to Bayer et al. [15]. The gels were stained with Coomassie brilliant blue R-250.

The concentration of the purified protein in solution was estimated spectrophotometrically ($\epsilon_{280} = 3.0$).

3. Results

3.1. Presence of biotin-binding proteins in Streptomyces cultures

Various strains of *Streptomyces* were grown under identical conditions. Following a 7-day growth period, the

Fig. 2. Comparative SDS-PAGE profiles of the streptavidin preparations from *S. avidinii* and *S. venezuelae*. Cultures of the desired strains were grown, the respective biotin-binding proteins were purified by affinity chromatography on iminobiotin columns and subjected to SDS-PAGE. Sa, streptavidin from *S. avidinii*; Sa v1, streptavidin v1 from *S. venezuelae* strain Tü 2460; Sa v2, streptavidin v2 from *S. venezuelae* strain Tü 2605; nSa, standard for native (full complement) streptavidin; cSa, standard for core (proteolysed) streptavidin. Note: in these particular batches, all of the proteins were degraded to some degree by resident proteinases in the culture medium. Compare with the native and core streptavidin standards (M_r 16 600 vs. 13 200, respectively). The experimental sample of streptavidin showed 4 bands; the smallest being slightly larger than core streptavidin. The native bands for streptavidins v1 and v2 both migrated at very slight, but consistently different rates than that of streptavidin, suggesting slight differences in their primary structures.
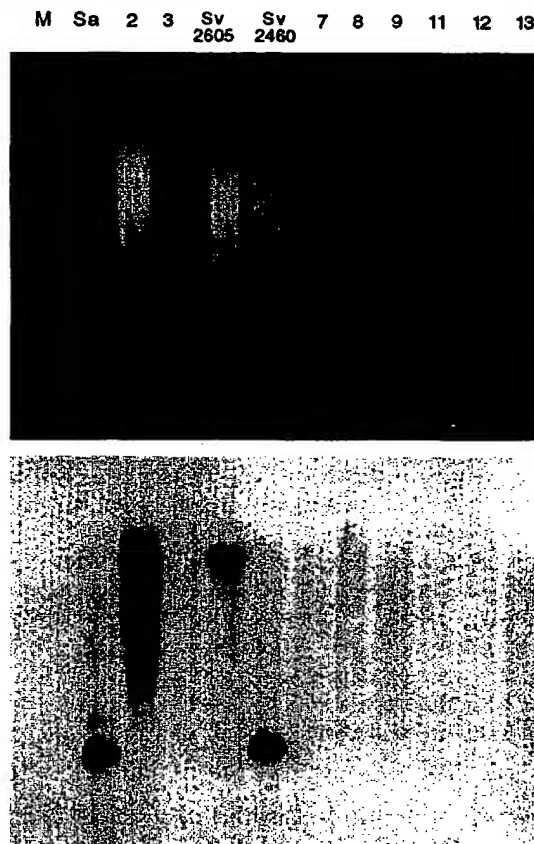
Fig. 3. Southern hybridization of genomic DNA from various species of *Streptomyces*. (Top) Agarose gel electrophoresis of *Bam*HI restriction digests of the indicated genomic DNAs. (Bottom) Southern hybridization using oligonucleotide probe based on the streptavidin gene of *S. avidinii*. Lanes: M, DNA markers, *S.a*, *S. avidinii*; 2, *S. lividans*; 3, *S. lactamdurans*; *S.v* 2605, *S. venezuelae* Tü 2605; *S.v* 2460, *S. venezuelae* Tü 2460; 7, *S. lavendulae*; 8, *S. jumonijinensis*; 9, *S. coelicolor*; 11 *S. griseus*; 12, *S. cattleya*; 13, *Streptomyces* aleph 50.

cell-free culture medium was examined for biotin-binding activity [17]. The results (Fig. 1) indicated that only *S. avidinii*, *S. venezuelae* Tü 2460 and 2605 exhibited such activity. In contrast, the type strain (ATCC 10712) of *S. venezuelae* did not contain such activity. Nor did the type strain (ATCC 8664) of *S. lavendulae*. Thus, only those cultures that were previously demonstrated to produce Acm-containing peptides showed biotin-binding activity.

The extent of biotin-binding activity in the three antibiotic-producing strains seemed quite similar, and we were interested in determining whether this activity reflected the presence of streptavidin-like biotin-binding proteins. We therefore subjected the cell-free growth medium from each strain to an affinity chromatographic procedure, designed to isolate streptavidin [18]. The purified proteins were examined for biotin-binding activity.

Following purification, the relative amounts of protein obtained from the cultures of both *S. venezuelae* strains, were found to be similar to that of *S. avidinii* (Table 1). In each case, SDS-PAGE analysis (Fig. 2) of the purified proteins revealed a high-molecular-weight band (ca. 17 kDa) for the monomer, which is considered to correspond to a native form of each protein as demonstrated previously for streptavidin [15]. In addition, the presence of one or more lower molecular-weight bands could be discerned, which reflect proteolytic breakdown products [19]. The mobilities of the bands in the *S. venezuelae* forms also differed slightly from that of *S. avidinii*.

In the case of the reference streptavidin sample prepared in this experiment, there appeared to be a higher level of endogenous proteinases in the culture, which led to a livelier degradation of the terminal appendages in the molecule. As documented earlier [15], the resultant low-

molecular-weight form is larger than core streptavidin by 4 amino acid residues, and this difference is detectable on SDS-PAGE gels.

3.2. Streptavidin-like genes in S. venezuelae Tü 2460 and 2605

The observed differences in the mobility profiles of the streptavidin-like proteins in the three strains suggested that the respective genes may also differ. We therefore synthesized two probes, based on the known sequence of the streptavidin gene. Genomic DNA from each of the strains was prepared and digested with a series of restriction enzymes. The Southern hybridization pattern of the *Bam*HI digests is shown in Fig. 3.

As can be seen from the figure, distinctive labeling was obtained for the digest of *S. avidinii* and the two *S. venezuelae* Tü strains. Thus, in accordance with the results observed for biotin-binding activity, the only strains which were labeled by the gene probes were the three strains which are known to express the Acm-based antibiotic. None of the other strains, including the *S. venezuelae* type strain (ATCC 10712), interacted with the probes.

Interestingly, the labeled band derived from *S. venezuelae* Tü 2460 was much closer in size to that of *S. avidinii* than to the labeled band from strain Tü 2605. This was true for all of the digests prepared using other restriction enzymes, e.g., *Kpn*I, *Sma*I and *Rsa*I as well (data not shown). It seems that the proximity of the gene from *S. venezuelae* strain Tü 2460 more closely resembles that of *S. avidinii* than that of strain Tü 2605. This observation was confirmed by cloning and sequencing the two genes (Fig. 4). Indeed, the sequences adjacent to the streptavidin

Fig. 4. Sequence alignment of the streptavidin genes from *S. avidinii* and *S. venezuelae*. Hyphens represent nucleotides in the respective genes for streptavidin v1 and v2 (Sa v1 and Sa v2) which are identical to those of the original streptavidin gene (Sa). Nucleotides that differ at a given position are indicated. Start and stop codons are marked by asterisks (* * *).
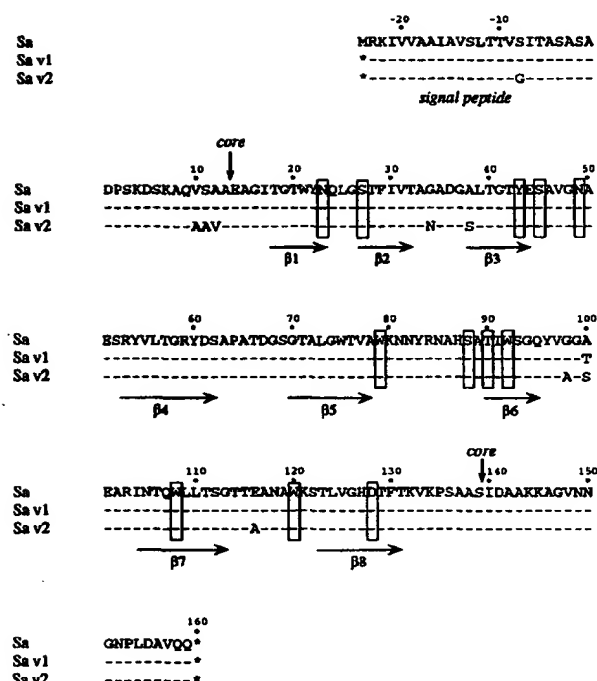
Fig. 5. Comparison of the deduced amino acid sequences of streptavidin v1 and v2 (Sa v1 and Sa v2) with the known primary structure of streptavidin (Sa). Amino acids known to contribute to the biotin-binding site in streptavidin are designated by boxes. The positions of the β structures (numbered) are shown by arrows. Residues which comprise the signal peptide are assigned negative numbers. Vertical arrows denote the N-and C-terminal proteolytic sites which delimit the stable streptavidin 'core'.

v1 gene (from *S. venezuelae* strain Tü 2460) were nearly identical to those of the original streptavidin gene. On the other hand, the sequences immediately upstream and downstream to the streptavidin v2 gene (from strain Tü 2605) showed many more nucleotide substitutions.

These comparative differences were commensurate within the corresponding genes themselves, i.e., the streptavidin v1 gene and the original streptavidin gene were much more similar between themselves than either was to streptavidin v2. Thus, streptavidins v1 and v2 showed 6 and 42 nucleotide substitutions, respectively, compared to the original streptavidin gene.

The deduced amino acid sequences of the streptavidins are presented in Fig. 5. Compared to the primary structure of streptavidin, streptavidin v1 showed only a single amino acid substitution at position 100 –a threonine instead of alanine. This residue is located within a loop which connects β-strands 6 and 7 of the β-barrel [20]. The residue is in an exposed position and is important neither to the structure of the protein nor to its biotin-binding function.

In streptavidin v2, 9 amino acid substitutions were apparent. One was in the signal peptide. Three more were in the extraneous N-terminal segment which is frequently

cleaved proteolytically during growth of the bacterium and postsecretory processing of the molecule. Consequently, within the reputed streptavidin v2 core protein (residues 14 to 139), only 5 amino acid substitutions were evident, compared to the known streptavidin core sequence. Like streptavidin v1, one of the substitutions occurred at position 100. Likewise, the other 4 substitutions are relatively conservative ones which occur in relatively unimportant residues located in exposed loops which interconnect β-strands. The residues of the β-strands per se are unaltered. All of the amino acid residues known to participate in the binding of biotin are likewise conserved. The only 'exciting' substitution is the replacement of alanine for glutamic acid at position 116, which would presumably result in a protein with a higher p*I*.

4. Discussion

In a sense, this work is a retrospective one. Three decades ago [7], a new synergistic antibiotic activity was described in a new species of *Streptomyces*, termed *S. avidinii*, which led to the discovery of the biotin-binding protein streptavidin [21]. Thus, streptavidin combines synergistically with the 'stravidins' [22] –namely, a group of di-or tri-peptides which contain an unusual amino acid, called amiclenomycin (Acm), which acts as a biotin antimetabolite [23,24].

Since these initial works, interest in this protein has assumed a practical nature, and the distinction of streptavidin has shifted to that of a preferred replacement for egg-white avidin in avidin-biotin technologies [25–27].

All but lost in the interim years was the fact that in the original work, similar antibiotic activity was also described in 11 *additional* strains [6]. All of these strains could be classified in a single, well-known species – *S. lavendulae*. The original authors, in their zeal to analyze the antibiotic activity from the newly classified *S. avidinii*, disregarded further analysis of the activities in the *S. lavendulae* strains. It is unfortunate that the latter strains are now unavailable. The question still stands as to the nature of putative streptavidin-like proteins in these strains.

For this reason, the more recent discovery [8] of the Acm antibiotic in two strains of *S. venezuelae* was intriguing to us. In particular, we were interested in re-addressing the question of whether new biotin-binding proteins are produced in species that express this type of antibiotic. If so, we wanted to isolate and characterize these proteins and their genes.

Our findings indeed indicate that the occurrence of streptavidin-like proteins –albeit unusual –may be more commonplace in *Streptomyces* than hitherto believed. Interestingly, the ATCC-derived 'type strains' of both *S. lavendulae* and *S. venezuelae* do not produce such biotin-binding proteins and their genomic DNAs fail to hybridize with streptavidin-specific probes. It thus appears that the

production of streptavidin-like proteins is not species-specific. Instead, the harboring of a streptavidin-like gene in an individual strain of *Streptomyces* and the expression of its product appear to be directly connected to its production of the synergistic Acm-based antibiotic activity.

Both traditional taxonomic methods [28] and ribosomal protein analysis [29] have indicated that *S. avidinii* and *S. lavendulae* are very closely related species. In fact, Ochi [30] suspects that the two species might be combined as a single taxon. On the other hand, *S. venezuelae* belongs to a different cluster group according to numerical classification [28] and appears to be more diverse than *S. avidinii* and *S. lavendulae*.

In this context, it is interesting that one of the streptavidins from *S. venezuelae* is more closely related to the original streptavidin than to its intraspecies cognate. In any case, the observed substitutions are certainly conservative ones and have little bearing on the chemical, structural and biotin-binding properties of the resultant protein molecule.

We have recently elucidated the three-dimensional crystal structure of egg-white avidin [4], and both its fold and the binding-site residues which are essential to binding biotin were compared with those of streptavidin [20,31]. When we initiated the present studies, we had hoped that the sequence of such streptavidin-like proteins would provide us with insight from nature into the types of mutations which would be worthwhile to perform on such a protein. Instead, nature seems to be telling us that there may be very little room for change in the amino acid sequences of these proteins in order to bind biotin.

The overall similarity in the sequences of egg-white avidin and core streptavidin is about 35%, but the similarity in the biotin-binding pocket is approx. 90%, indicating the importance of selected residues to the binding of biotin. The consequence of the differences may be the observed reduction of two orders of magnitude in the binding affinity, in favor of the egg-white protein [32]. The necessity of certain binding-site residues was also implied in our recent studies on antibiotin antibodies [5], in which binding-site motifs similar to those of the binding sites of avidin and streptavidin were found, although the binding affinity of the antibody for biotin is significantly lower. Markedly reduced biotin-binding activities were also demonstrated recently on peptide segments isolated from avidin [33] and in a library of peptides synthesized by bacteriophages [34]. The logical conclusion from these studies is that mutations in the binding site residues of avidin or streptavidin will lead to a weaker binding of biotin. These studies also suggest that the strong binding to biotin may play a major biological function, although its exact role in nature and its implications to the antibiotic activity (other than complete inactivation thereof) remains a mystery.

In conclusion, the streptavidins from *S. avidinii* and *S. venezuelae* are very similar molecules. They bind biotin similarly and to the same extent, and are expected to

display the same type of three-dimensional fold. Since streptavidins are produced by several related species of *Streptomyces*, the original classification of *S. avidinii* as a separate species should be reevaluated. The discovery, cloning and sequencing of additional streptavidin-like genes together with site-directed mutagenesis studies should shed further light on the status of such biotin-binding proteins.

Acknowledgements

References

[1] Gitlin, G., Bayer, E.A. and Wilchek, M. (1988) Biochem. J. 256, 279–282.

[2] Gitlin, G., Bayer, E.A. and Wilchek, M. (1988) Biochem. J. 250, 291–294.

[3] Gitlin, G., Bayer, E.A. and Wilchek, M. (1990) Biochem. J. 269, 527–530.

[4] Livnah, O., Bayer, E.A., Wilchek, M. and Sussman, J.L. (1993) Proc. Natl. Acad. Sci. USA 90, 5076–5080.

[5] Bagci, H., Kohen, F., Kuscuoglu, U., Bayer, E.A. and Wilchek, M. (1993) FEBS Lett. 322, 47–50.

[6] Stapley, E.O., Mata, J.M., Miller, I.M., Demny, T.C. and Woodruff, H.B. (1963) Antimicrob. Agents Chemother. 3, 20–27.

[7] Chaiet, L., Miller, T.W., Tausig, F. and Wolf, F.J. (1963) Antimicrob. Agents Chemother. 3, 28–32.

[8] Kern, A., Kabatek, U., Jung, G., Werner, R.G., Poetsch, M. and Zähner, H. (1985) Liebigs Ann. Chem. 1985, 877–892.

[9] Shiffman, D., Mevarech, M., Jensen, S.E., Cohen, G. and Aharonowitz, Y. (1988) Mol. Gen. Genet. 214, 562–569.

[10] Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.

[11] Argaraña, C.E., Kuntz, I.D., Birken, S., Axel, R. and Cantor, C.R. (1986) Nucleic Acids Res. 14, 1871–1882.

[12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

[13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.

[14] Bayer, E.A., Ben-Hur, H. and Wilchek, M. (1990) Methods Enzymol. 184, 217–223.

[15] Bayer, E.A., Ben-Hur, H. and Wilchek, M. (1990) Methods Enzymol. 184, 80–89.

[16] Green, N.M. (1970) Methods Enzymol. 18, 418–424.

[17] Bayer, E.A., Ben-Hur, H. and Wilchek, M. (1986) Anal. Biochem. 154, 367–370.

[18] Bayer, E.A., Ben-Hur, H., Gitlin, G. and Wilchek, M. (1986) J. Biochem. Biophys. Methods 13, 103–112.

[19] Bayer, E.A., Ben-Hur, H., Hiller, Y. and Wilchek, M. (1989) Biochem. J. 259, 369–376.

[20] Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J. and Salemme, F.R. (1989) Science 243, 85–88.
[21] Chaiet, L. and Wolf, F.J. (1964) Arch. Biochem. Biophys. 106, 1–5.
[22] Baggaley, K.H., Blessington, B., Falshaw, C.P., Ollis, W.D., Chaiet, L. and Wolf, F.J. (1969) Chem. Commun. 1969, 101–102.
[23] Okami, Y., Kitahara, T., Hamada, M., Naganawa, H., Kondo, S., Maeda, K., Takeuchi, T. and Umezawa, H. (1974) J. Antibiotics 27, 656–664.
[24] Hotta, K., Kitahara, T. and Okami, Y. (1975) J. Antibiotics 28, 222–228.
[25] Bayer, E.A. and Wilchek, M. (1994) in Egg Uses and Processing Technologies (Sim, J.S. and Nakai, S., eds.), pp. 158–176, CAB International, Wallingford.
[26] Wilchek, M. and Bayer, E.A. (1988) Anal. Biochem. 171, 1–32.
[27] Wilchek, M. and Bayer, E.A. (1990) Methods Enzymol. 184, 746.
[28] Williams, S.T., Goodfellow, M., Alderson, G., Wellington, E.M.H., Sneath, P.H.A. and Sackin, M.J. (1983) J. Gen. Microbiol. 129, 1743–1813.
[29] Ochi, K. (1989) J. Gen. Microbiol. 135, 2635–2642.
[30] Ochi, K. (1989) Actinomycetologica 3, 111–114.
[31] Hendrickson, W.A., Pähler, A., Smith, J.L., Satow, Y., Merritt, E.A. and Phizackerley, R.P. (1989) Proc. Natl. Acad. Sci. USA 86, 2190–2194.
[32] Green, N.M. (1990) Methods Enzymol. 184, 51–67.
[33] Hiller, Y., Bayer, E.A. and Wilchek, M. (1991) Biochem. J. 278, 573–585.
[34] Saggio, I. and Laufer, R. (1993) Biochem. J. 293, 613–616.

Human Serum Biotinidase

cDNA CLONING, SEQUENCE, AND CHARACTERIZATION*

Heath Cole‡, Thomas R. Reynolds§, Jean M. Lockyer¶, Gregory A. Buck§, Ted Denson‡, J. Edward Spence‡|, Jeanne Hymes‡, and Barry Wolf‡‡‡**

From the Departments of ‡Human Genetics, §Microbiology and Immunology, and |Pediatrics, Medical College of Virginia / Virginia Commonwealth University, Richmond, Virginia 23298 and the ¶Department of Medicine, Tulane University School of Medicine, New Orleans, Louisiana 70112

Biotinidase (EC 3.5.1.12) catalyzes the hydrolysis of biocytin, the product of biotin-dependent carboxylase degradation, to biotin and lysine. Biotinidase deficiency is an inherited metabolic disorder of biotin recycling that is characterized by neurological and cutaneous abnormalities, and can be successfully treated with biotin supplementation. Sequences of tryptic peptides of the purified human serum enzyme were used to design oligonucleotide primers for polymerase chain reaction amplification from human hepatic total RNA to generate putative biotinidase cDNA fragments. Sequence analysis of a cDNA isolated from a human liver library by plaque hybridization with the largest cDNA probe revealed an open reading frame of 1629 bases encoding a protein of 543 amino acid residues, including 41 amino acids of a potential signal peptide. Comparison of the open reading frame with the known biotinidase tryptic peptides and recognition of the expressed protein encoded by this cDNA by monoclonal antibodies prepared against purified biotinidase demonstrated the identity of this cDNA. Southern analyses suggested that biotinidase is a single copy gene and revealed that human cDNA probes hybridized to genomic DNA from mammals, but not from chicken or yeast. Northern analysis indicated the presence of biotinidase mRNA in human heart, brain, placenta, liver, lung, skeletal muscle, kidney, and pancreas.

Biotinidase (EC 3.5.1.12) catalyzes the release of biotin, an essential B-complex vitamin, from biocytin, the degradative product of the four biotin-dependent holocarboxylases, pyruvate carboxylase, acetyl-CoA carboxylase, propionyl-CoA carboxylase and β-methylcrotonyl-CoA carboxylase, and from proteolytically degraded dietary proteins (1). Mammals cannot synthesize biotin and, therefore, must obtain the vitamin from their diet and from recycling endogenous biotin.

Biotinidase deficiency, an autosomal recessive disorder, results in a secondary biotin deficiency that leads to multiple carboxylase deficiency (2). Clinical features of untreated individuals with profound biotinidase deficiency (<10% of mean normal activity) include seizures, hypotonia, skin rash, alopecia, developmental delay, conjunctivitis, visual problems, hearing loss, metabolic ketolactic acidosis, organic acidemia, and hyperammonemia (1). There is variability in the expression of these features and in the age of onset of symptoms, even within the same family (3). Because biotin therapy, when initiated early, prevents many clinical and biochemical symptoms of this disorder (1), newborn screening for biotinidase deficiency has been implemented in many states in the United States and in many countries (4).

We now describe the cloning and sequencing of the cDNA encoding normal human biotinidase. The tissue distribution of the biotinidase mRNA, copy number, and preliminary analysis of conservation of the genomic gene for biotinidase were determined.

EXPERIMENTAL PROCEDURES

General Methods—Standard methods were performed (5), except where noted. Restriction endonucleases and DNA modifying enzymes were obtained from Life Technologies, Inc.. Membranes were prehybridized in 5 × SSC, 5 × Denhardt's reagent (0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin), 2% SDS, and 100 µg/ml sonicated salmon sperm DNA at 60 °C for 4 h. Hybridizations were performed in fresh prehybridization solution with 10% dextran sulfate at 60 °C for 14–18 h. The final stringency of washing conditions was 0.2 × SSC, 1% SDS at 60 °C for 30 min, except where noted. Filters were dried, covered in plastic wrap and exposed to X-Omat AR scientific imaging film (Kodak) with Cronex intensifying screens (DuPont) at –70 °C for 24 h, unless otherwise noted. Oligonucleotides used in polymerase chain reactions (PCR)[1] and sequencing procedures were synthesized by the Louisiana State University Core Facility (New Orleans, LA) and the Nucleic Acid Synthesis and Analysis Laboratory (Medical College of Virginia/Virginia Commonwealth University (MCV/VCU), Richmond, VA). Radiolabeled deoxynucleotides were obtained from DuPont NEN (Boston, MA).

cDNA Probes—BTD400 was liberated from pCR1000 upon digestion with *Eco*RI and *Hind*III. BTD2000 was excised from pBluescript SK with *Bam*HI and *Apa*I. Inserts were isolated from a 1.3% low melting point agarose gel, purified by extraction with phenol:chloroform and recovered by ethanol precipitation. The β-actin cDNA was obtained from Clontech (Palo Alto, CA). cDNAs were radiolabeled with [α-³²P]dCTP (3000 Ci/mmol) (DuPont NEN) using an oligolabeling kit (Pharmacia LKB Biotechnology Inc.).

Preparation of Tryptic Peptides of Biotinidase—Enzymatically active biotinidase was purified 22,000-fold from pooled normal human serum (6). A single amino acid N terminus was found on analysis of the protein. A single silver-staining protein was observed on SDS and native polyacrylamide gel electrophoresis (PAGE). Polyclonal and monoclonal antibodies prepared against the purified serum enzyme detected a single, approximately 74-kDa protein in serum on immunoblots of SDS-PAGE (7). Using these antibodies, we have identified several individuals with profound biotinidase deficiency who lack cross-reacting material to an-

[1] The abbreviations used are: PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); kb, kilobase(s); Tricine, *N*-tris(hydroxymethyl)methylglycine.

tibodies prepared against biotinidase in their serum on SDS-PAGE and isoelectric focusing electrophoresis (7). Purified biotinidase was treated with trypsin, the resulting peptides were separated by high performance liquid chromatography, and 11 were sequenced by Edman degradation (8) at the Yale University Protein and Nucleic Acid Chemistry Facility (W. M. Keck Foundation Resource Laboratory, New Haven, CT).

PCR Conditions and Generation of cDNA Probes—Three sets of oligonucleotide primers were derived from the amino acid sequence of several tryptic peptides of the biotinidase protein (Fig. 1) and from gene-specific DNA sequence. The primers in set 1 are degenerate (1A, sense, 32-fold degeneracy; 1B, antisense, 96-fold degeneracy), whereas those in the second set (2A, sense; 2B, antisense) were generated according to codon usage frequencies (9). The third set consists of a gene-specific oligonucleotide (3B, antisense, 5'-GCCCCTGATGGCCATA-CAACT-3') synthesized from the reverse complement of the DNA sequence of BTD300 and a degenerative primer pool designed from the N-terminal peptide in which inosine was substituted in positions where degeneracy indicated either an adenosine or guanine (3A, sense, 8192-fold degeneracy).

RNA was isolated from frozen human liver by the guanidinium thiocyanate cesium chloride density gradient centrifugation method (5). Reverse transcription was performed in a final volume of 50 μl with 10 μg of total RNA, 100 ng of primer 1B, 50 mM Tris-HCl (pH 8.3), 150 mM KCl, 10 mM $MgCl_2$, 0.5 mM of each dNTP, 0.5 μg of actinomycin D, 10 mM dithiothreitol, 20 mM β-mercaptoethanol, and 15 units of avian myeloblastosis virus reverse transcriptase. Following incubation at 42 °C for 60 min, the RNA was precipitated with ethanol, resuspended in 20 μl of 5 × PCR buffer (50 mM Tris-HCl (pH 9.0), 50 mM KCl, 15 mM $(NH_4)_2SO_4$, and 7 mM $MgCl_2$). The RNA was used as a template in PCR, performed in a final volume of 100 μl, with 0.1 μg of primers 1A and 1B, 0.5 mM of each dNTP and 5 units of *Taq* polymerase (BioRes). After amplification in a thermocycler (NOLA Scientific, New Orleans, LA) was performed for 40 cycles (92 °C for 1.5 min, 55 °C for 3 min, 72 °C for 4 min), the DNA was precipitated with ethanol and used as a template for an additional PCR amplification using 0.2 μg of primers 2A and 2B for 20 cycles (92 °C for 1.5 min, 52 °C for 3 min, 72 °C for 2.5 min). The resulting 300-bp fragment (BTD300) was purified by electroelution from a 1.5% agarose gel and subcloned into pBluescript KS. Dideoxynucleotide sequencing (10) was performed with T3 and T7 universal oligonucleotide primers (Stratagene, La Jolla, CA).

A second reverse transcription reaction was performed with 2 μg of total RNA, isolated from fresh human liver obtained from the MCV/VCU Human Tissue Acquisition and Histopathology Facility, and 5 ng of oligo(dT)$_{15}$ primer (Promega, Madison, WI) as described (5). One-half μg of primers 3A and 3B and 2.5 units of AmpliTaq polymerase (Perkin-Elmer Corp.) were used in PCR for 45 cycles (94 °C for 1 min, 55 °C for 1 min, 72 °C for 3 min) with an initial denaturation step at 94 °C for 5 min and a final extension period of 5 min at 72 °C in a Twin Block System PCR thermocycler (Ericomp Inc., San Diego, CA). The resulting 400-bp putative biotinidase cDNA fragment (BTD400) was directly purified by high performance liquid chromatography on an analytical (4 × 250 mm) Nucleo Pac PA-100 anion exchange column (Dionex, Sunnyvale, CA) at a flow rate of 1.5 ml/min. The product was fractionated using a gradient starting at 95% buffer A (20 mM Tris-HCl, pH 7.3) and

5% buffer B (20 mM Tris-HCl, 1.0 M NaCl at pH 7.3) increasing linearly to 50% buffer B at 10 min and subsequently increasing to 100% buffer B at 40 min. The product eluted at 23 min and was collected, desalted using a Centricon-30 (Amicon, Beverly, MA), and lyophilized. BTD400 was cloned into the pCR1000 vector according to the manufacturer's instructions (TA cloning system, Invitrogen, San Diego, CA). Both strands of BTD400 were sequenced using Sequenase 2.0 DNA polymerase (U. S. Biochemical Corp.) and fluorescently labeled terminator nucleotides (DuPont) according to the manufacturer's instructions (11). The sequence was analyzed using the GENESIS 2000 system (DuPont).

Isolation and Sequencing of Biotinidase cDNA Clone—A human liver cDNA library cloned in the Uni-ZAP XR vector was obtained from Stratagene. Approximately 5 × 10^5 recombinant bacteriophage were screened by plaque hybridization with the BTD400 cDNA probe under conditions described under "General Methods." Bacteriophage DNAs containing putative biotinidase inserts were isolated (λ Quick!, BIO101, La Jolla, CA). 0.5–1.0 μg of the DNA was digested with HindIII, separated on a 0.8% agarose gel, and identified by Southern analysis with the BTD400 cDNA probe, as described previously. Recovery of phagemids (pBluescript SK) from putative purified bacteriophage particles was performed in the presence of R408 helper phage according to the *in vivo* excision protocol (Stratagene).

Plasmid DNAs containing putative biotinidase cDNA inserts were purified by cesium chloride density gradient centrifugation (5). Sequencing reactions were performed with 1 μg of plasmid DNA, 3.2 pmol of primer, fluorescently labeled terminator nucleotides (DyeDeoxy™) and AmpliTaq DNA polymerase according to manufacturer's instructions (Applied Biosystems, Foster City, CA) in a Perkin-Elmer model 9600 thermocycler. DNA sequence was determined for both strands using a model 373A Automated DNA Sequencer (Applied Biosystems). Sequence editing was performed using the Genetics Computer Group Package (UW Biotechnology Center, Madison, WI) (12).

Expression of the Cloned Biotinidase cDNA—In order to perform expression studies, the Bluescript plasmid containing the biotinidase cDNA was modified so that the cloned insert was in the same reading frame as that of β-galactosidase. Plasmid DNA (3.6 μg) was linearized with BamHI, blunt ends were created by the incorporation of dNTPs (4 mM each) with the Klenow fragment of DNA polymerase, and the plasmid was religated with T4 bacteriophage DNA ligase. Plasmid DNA, isolated by cesium chloride density gradient centrifugation, was sequenced from the M13R primer in the sense orientation with a dideoxynucleotide sequencing kit (Stratagene) and [α-^{35}S]dATP (500 Ci/mmol). The modified plasmid and a control Bluescript plasmid (containing no insert) were each transformed into XL1Blue competent cells and induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside for 2.5 h. The cells were sedimented by centrifugation, washed three times in phosphate-buffered saline (pH 7.4), and then subjected to a French pressure cell (SLM Aminco, Urbana, IL). The extract was centrifuged (1000 × *g*), and 20 μl of the supernatant were electrophoresed in a 12% SDS-PAGE Tricine gel (Millipore Corp., Bedford, MA). The gel was then electroblotted onto nitrocellulose (Bio-Rad) as described (13). The membrane was pretreated with nonfat dry milk and then incubated with an IgG preparation of monoclonal antibodies against purified human serum biotinidase (Hybridoma-Monoclonal Antibody Laboratory, MCV/VCU)
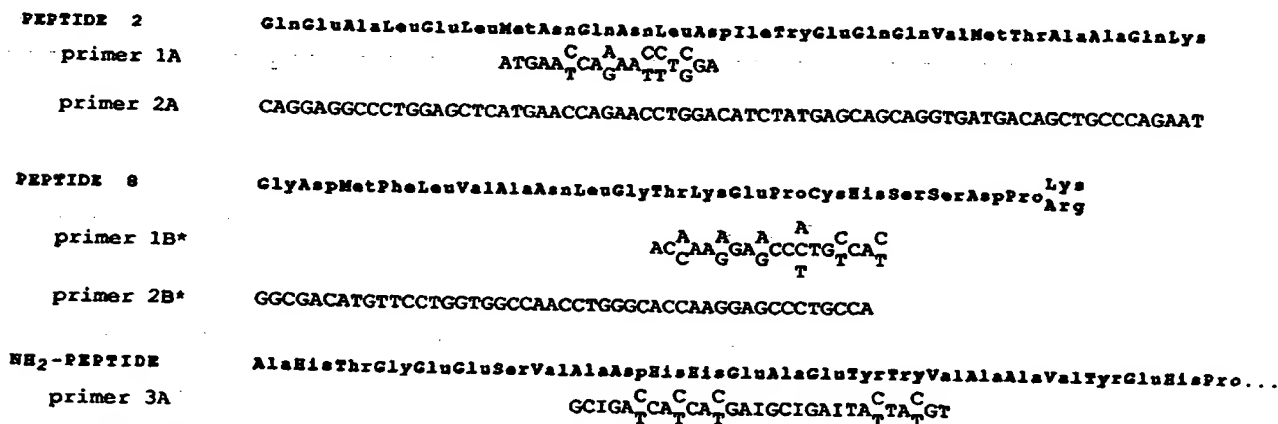

FIG. 1. **Coding nucleotide sequence of oligonucleotide primers (5' → 3') used in PCR to generate a screening probe.** *Asterisk* indicates that the reverse complement was used for the primer sequence.

as described by Hart *et al.* (7). Immunoreactive species were detected using the enhanced chemiluminescence (ECL) system according to the manufacturer's instructions (Amersham Corp.).

Genomic Southern Analyses—Total genomic DNA was isolated from lymphocytes of an individual with normal serum biotinidase activity using a model 340A nucleic acid extractor (Applied Biosystems). Ten µg of DNA were digested with *Bam*HI, *Bgl*I, *Eco*RI, *Eco*RV, *Xba*I, or *Xho*I. After electrophoresis through a 0.8% agarose gel, the DNA was transferred onto a 0.45-µm Nytran nylon membrane (Schleicher & Schuell) as described by Southern (14). A ZOO-blot containing 8 µg of *Eco*RI-digested genomic DNA from nine eukaryotic species was purchased from Clontech. These membranes were hybridized with radiolabeled BTD2000 and exposed to autoradiography for 5 and 14 days, respectively, as described under "General Methods." After the BTD2000 probe was removed from the ZOO-blot with 0.2 M NaOH at 60 °C for 20 min, the filter was hybridized with the human BTD400 cDNA probe, washed at conditions of lower stringency (twice in 2 × SSC at 55 °C for 30 min, then once in 0.1 × SSC at room temperature for 5 min), and exposed to x-ray film for 5 days.

RESULTS AND DISCUSSION

Cloning and Sequence of cDNA Encoding Biotinidase—The amino acid sequences of 11 tryptic peptides, including the N-terminal peptide, were determined by Edman degradation. No match for any of these sequences was found in the Protein Identification Resource, GenBank, or Swiss Protein data bases. A reverse transcription polymerase chain reaction was performed using the degenerate oligonucleotide primers (1A and 1B) derived from the amino acid sequences of peptide 2 and peptide 8, respectively (Fig. 1), and then using a set of primers (2A and 2B) generated according to codon usage frequencies. The resulting 300-bp putative biotinidase fragment (BTD300) was determined to contain 186 bp that were unique and not included in the primer sequences. Identity between five peptides (3, 4, 5, 6, and 7) and the amino acid sequences deduced from the nucleotide sequence of BTD300 indicated that

FIG. 2. **Nucleotide and deduced amino acid sequences of human serum biotinidase.** Two potential ATG initiation codons are *double underlined* (numbered as bases 1 and 61) and could encode signal peptides of 41 and 21 amino acids, respectively. An Ala residue and a Gly residue are located at amino acid positions −3 and −1, respectively, to the N terminus of the mature protein, which is consistent with a signal peptidase cleavage site (22). An open reading frame of 1629 bp, relative to the first ATG codon and the termination codon (indicated by ***) was present. The N terminus of the mature serum protein (starting at base 124) is indicated in *uppercase letters*. The amino acid sequences of 11 tryptic peptides from purified human serum biotinidase are *underlined*. Six potential N-linked glycosylation sites (Asn-X-Thr/Ser) are noted with *asterisks* above the sequence. The cDNA clone also includes 35 bp 5′ of the first ATG codon and 332 bp of 3′-untranslated sequence, with a polyadenylation signal (AATAAA) at base 1932 (*boldface*) located 24 bp upstream from the 20-bp poly(A) tract.

BTD300 represented a region of the biotinidase gene. In order to obtain a larger portion of the biotinidase cDNA for use as a probe, an additional reverse transcription polymerase chain reaction, performed using a gene-specific oligonucleotide (3B) synthesized from the DNA sequence of BTD300 and a degenerative primer pool (3A) designed from the N-terminal peptide, yielded a 400-bp putative biotinidase fragment (BTD400). The nucleotide sequence of BTD400 matched that of BTD300 and showed sequence identity with the amino acid sequence of five tryptic peptides (2, 3, 4, 5, and 6), thus confirming that BTD400 is a portion of the biotinidase cDNA.

One-fourth of the 2×10^6 primary plaques in a bacteriophage human liver cDNA library was screened by plaque hybridization with radiolabeled BTD400. A liver library was chosen because the liver appears to be the primary source of serum biotinidase (15–17). Two clones that produced strong hybridization signals with BTD400 on duplicate filters in the primary screen were isolated after a tertiary screen. Hybridization between the BTD400 probe and HindIII-digested bacteriophage DNA isolated from these two clones indicated that the inserts contained portions of the biotinidase cDNA. Both strands of the largest insert were sequenced with universal primers M13R, T7, and M13F and gene-specific oligonucleotides targeted to the newly obtained biotinidase sequence.

The complete nucleotide sequence of the longest cDNA insert (2016 bp) and the deduced amino acid sequence are shown in Fig. 2. All peptides obtained from the tryptic digestion of the biotinidase protein were identified in the sequence encoded by the cDNA. There were no peptide sequences that were not contained in this sequence. No matches or significant homologies with either the DNA or amino acid sequences of biotinidase were found in searches of GenBank, EMBL, VecBase, PIR-Nucleic, and PIR-Protein data bases. This cDNA encodes for a mature protein of 502 amino acids with a molecular mass of 56,771 Da, which is similar to the 60 kDa reported for the glycanase-treated serum enzyme (13). The amino acid composition of the encoded mature enzyme also compares favorably with previously published analyses (18, 19) and with our unpublished data. Glycosylation of the protein would result in an increase in molecular mass of 14–23 kDa, given that the glycosylation structures range in mass from approximately 2.4 kDa for sialated biantennary to 3.8 kDa for sialated tetraantennary structures with fucose residues (20), and assuming that all six of the potential glycosylation sites in the protein are glycosylated. The molecular mass of the glycosylated enzyme is estimated to be between 70.8 and 79.8 kDa, which is consistent with that of the glycosylated serum enzyme reported by our laboratory and others (6, 18, 19, 21).

Expression analysis of the cDNA insert was performed to demonstrate that the cloned cDNA encoded biotinidase. The cDNA insert was modified into the same reading frame as β-galactosidase as described under "Experimental Procedures." Sequencing of the plasmid confirmed that the BamHI site was absent and that the cloned insert was in the same reading frame as β-galactosidase. Expression of this cloned cDNA yielded a protein with an estimated molecular mass of 53 kDa that reacted immunologically with anti-biotinidase, thus showing that this cDNA encodes biotinidase.

Expressions, Copy Number, and Conservation of the Biotinidase Gene—Northern analysis using the radiolabeled BTD2000 cDNA probe revealed a 2.0-kb hybridization signal in multiple human tissues (Fig. 3). Hybridization of the BTD2000 probe with genomic DNA digested with BamHI, EcoRI, EcoRV, and XhoI produced a single hybridization signal at 7.5, 11.6, 13.2, and greater than 24.0 kb, respectively (results not shown). Because these enzymes do not cleave the biotinidase cDNA, these results suggest that biotinidase is a single copy gene with
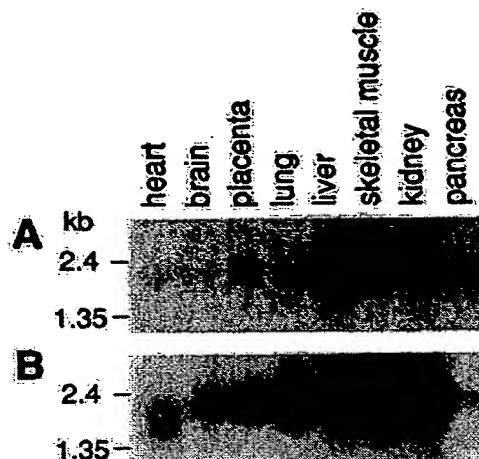

FIG. 3. **Northern analysis of poly(A⁺) RNA from multiple human tissues.** A Northern blot of poly(A⁺) RNA (2 μg) isolated from multiple human tissues was purchased from Clontech. A, after hybridization with the BTD2000 cDNA probe and a 14-day exposure to autoradiography, the full-length biotinidase mRNA (2.0 kb) was observed in mRNA isolated from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. Previous studies have found detectable biotinidase activities in mammalian (non-human) lung, liver, skeletal muscle, kidney, pancreas, and heart (15), but little or no biotinidase activity in mammalian brain (15, 23, 24). After removing the BTD2000 probe by boiling the blot for 10 min, the filter was exposed to autoradiography for 2 days to ensure that the probe was removed. B, the integrity of the mRNAs in all lanes was demonstrated by the presence of the expected 2.0-kb hybridization of the radiolabeled β-actin cDNA probe after exposure to x-ray film for 14 h. The additional 1.8-kb signal present in the lanes containing the heart and skeletal muscle mRNAs is due to an isoform of β-actin found in these tissues (25).

no indication of pseudogenes.

Hybridization was observed between human cDNA probes, BTD400 and BTD2000, and the digested DNA from human, monkey (Rhesus), rat (Sprague-Dawley), mouse (BALB/c), dog, and cow (data not shown). The hybridization pattern of the BTD400 cDNA probe with these mammalian DNAs was identical to that observed between the BTD2000 cDNA probe. Hybridization did occur between the BTD400 probe and DNA from rabbit, but not between the BTD2000 probe and rabbit DNA, which may be attributed to the lower stringency washing conditions used in the experiment with the BTD400 probe. Both human cDNA probes did not hybridize to DNA from chicken or yeast (*Saccharomyces cerevisiae*) under the experimental conditions used. Using a method that measures the ability of biotinidase to hydrolyze N-(d-biotinyl)-p-aminobenzoate, an analogue of biocytin, biotinidase activity has been detected in human, rat, mouse, dog, cow, and chicken by Pispa (15). Failure of the human probes to hybridize to chicken DNA in our analysis may indicate that there is insufficient homology between the human cDNA probes and the genomic chicken DNA sequences to allow hybridization or that the conditions used in the analysis were too stringent for hybridization to occur.

Characterization of mutant biotinidase genes are in progress and should provide insight into the genotype/phenotype relationships in biotinidase deficiency and a better understanding of the clinical expression and variability of the disorder. In addition, determination of the mutations and the expression of these genes should provide insight about the protein domains that are important for enzyme function and the role of biotinidase in normal metabolism and nutrition.

vided by Brian Barnstein, Christain Lytle, Alison Shakarian, Tom Turner, and Rollin Wright.

REFERENCES

1. Wolf, B. (1992) in *The Metabolic Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) pp. 2083–2103, McGraw-Hill, Inc., New York
2. Wolf, B., Grier, R. E., Allen, R. J., Goodman, S. I., and Kien, C. L. (1983) *Clin. Chim. Acta* **131**, 273–281
3. Wolf, B., Grier, R. E., Allen, R. J., Goodman, S. I., Kien, C. L., Parker, W. D., Howell, D. M., and Hurst, D. L. (1983) *J. Pediatr.* **103**, 233–237
4. Wolf, B., Heard, G. S., Jefferson, L. G., Weissbecker, K. A., Secor McVoy, Jr., Nance, W. E., Mitchell, P. L., Lambert, F. W., and Linyear, A. S. (1986) *J. Inherited Metab. Dis.* **9**, Suppl. 2, 303–306
5. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
6. Wolf, B., Miller, J. B., Hymes, J., McVoy, J. S., Ishikawa, Y., and Shapira, E. (1987) *Clin. Chim. Acta* **164**, 27–32
7. Hart, P. S., Hymes, J., and Wolf, B. (1992) *Am. J. Hum. Genet.* **50**, 126–136
8. Stone, K. L., LoPresti, M. B., Crawford, J. M., DeAngelis, R., and Williams, K. R. (1989) in *A Practical Guide to Protein and Peptide Purification of Microsequencing* (Matsudaria, P., ed) pp. 31–47, Academic Press, Orlando, FL
9. Lathe, R. (1985) *J. Mol. Biol.* **183**, 1–12
10. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
11. Genesis 2000 Application Brief 30 (AB30-1) (1991) in *Molecular Genetics Customer Support Group*, DuPont, Wilmington, DE
12. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
13. Hart, P. S., Hymes, J., and Wolf, B. (1991) *Clin. Chim. Acta* **197**, 257–264
14. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517
15. Pispa, J. (1965) *Ann. Med. Exp. Biol. Fenn.* **43**, (suppl.) 1–39
16. Weiner, D. L., Grier, R. E., Watkins, P., Heard, G. S., and Wolf, B. (1983) *Am. J. Hum. Genet.* **34**, 56A
17. Grier, R. E., Heard, G. S., Watkins, P., and Wolf, B. (1989) *Clin. Chim. Acta* **186**, 397–400
18. Chauhan, J., and Dakshinamurti, K. (1986) *J. Biol. Chem.* **261**, 4268–4275
19. Oizumi, J., Hayakawa, K., and Hosoya, M. (1989) *Biochimie* **71**, 1163–1169
20. Lee, K. B., Loganathan, D., Merchant, Z. M., and Linhardt, R. J. (1990) *Appl. Biochem. Biotechnol.* **23**, 53–80
21. Craft, D. V., Goss, N. H., Chandramouli, N., and Wood, H. G. (1985) *Biochemistry* **24**, 2471–2476
22. Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391–409
23. Suchy, S. F., Secor, McVoy, Jr., and Wolf, B. (1985) *Neurology* **35**, 1510–1511
24. Oizumi, J., and Hayakawa, K. (1990) *Arch. Biochem. Biophys.* **278**, 381–385
25. Pari, G., Jardine, K., and McBurney, N. W. (1991) *Mol. Cell. Biol.* **11**, 4796–4803

BPABA are cleaved at much greater rates than biocytin. Biotinidase from S. faecalis can cleave biotinyl-β-alanine and biocytin at approximately equal rates.

Inhibition

Biotin is a competitive inhibitor of biotinidase at acidic pH when either biocytin or BPABA is the substrate. Biotin is not inhibitory with BPABA as the substrate at pH 7.4,[11] and the K_i for biotin is increased markedly with biocytin as substrate at alkaline pH.[17] Biotinidase activity is not competitively inhibited by dl-dethiobiotin, l-biotin, l-lysine, ε-aminoca-proic acid, fatty acids, d-norbiotin, d-biotin sulfone, or PABA.[11,17]

Sulfhydryl inhibitors inactivate biotinidase. Hog liver and serum biotinidases are completely inactivated by 0.1 mM p-chloromercuriben-zoate,[11] and the enzymes in human plasma[1] and hog kidney[15] are totally inactivated by 10 μM p-chloromercuribenzoate.[11] Monoiodoacetate also inhibits enzyme activity, whereas N-ethylmaleimide or arsenite does not inhibit.[11] Partially purified biotinidases from L. casei and S. faecalis are not inhibited by 0.1 mM p-chloromercuribenzoate.[9,15]

Evidence for the presence of a serine hydroxyl group in or near the active site of biotinidase is equivocal. Hog kidney biotinidase has only 15% of normal activity when treated with 1 μM diisopropyl fluorophos-phate,[15] and the activity of the human plasma enzyme is inhibited 99% by 0.1 μM phenylmethylsulfonyl fluoride.[1] Phenylmethylsulfonyl fluoride does not inhibit human serum biotinidase at 1 mM.[17] Diisopropyl fluoro-phosphate at 1 mM does not completely inactivate the hog serum enzyme and inhibits hog liver enzyme only about 30%.[11] Partially purified biotini-dase from L. casei is inhibited 50% by 0.4 mM diisopropyl fluorophos-phate.[15]

The K_i for biotin using BPABA as substrate is 100 μM at pH 5.0 and 1.3 mM at pH 6.5 for hog serum biotinidase.[11] The K_i for the hog liver enzyme is 40 μM at pH 5.0 and 0.9 mM at pH 6.5.[11] The K_i for the human serum enzyme is 87 μM at pH 5.5 and 1.3 mM at pH 7.[17] The K_i for biotin for human serum enzyme when biocytin is the substrate is 20 μM at pH 5.5 and 0.5 mM at pH 7.[17] The lack of inhibition by biotin at pH 7.4 is thought to be due to slower dissociation of biotin from the enzyme. The slow turnover of product results in less available, inhibitable en-zyme.[11,17]

The effect of pH on the inhibition of biotinidase by biotin appears not to be due to changes in the affinity of the enzyme for the substrate. The K_m values for BPABA and biocytin are virtually unchanged between pH 5.5 and 7.0.[11,17]

Stability

Biotinidase loses activity during freezing and thawing but is stable when stored at −70°.[1,17] After ammonium sulfate precipitation the human plasma biotinidase loses 67% of its initial activity in 9 months.[1] Dialyzing hog serum biotinidase at between pH 5.0 and 6.0 stabilizes the enzyme.[11] Buffers with a salt concentration of at least 0.1 M prevent loss of activ-ity.[17] Human serum biotinidase can be stored at 4° in 0.1 M phosphate buffer (pH 6.0) containing 1 mM 2-mercaptoethanol and 1 mM EDTA for 1 month without loss of activity.[17]

The addition of mercaptoethanol (0.1–0.2 mM) or dithiothreitol (1–2 mM) appears to protect or stabilize the enzyme.[1,11,17,32] Heavy metals, such as copper and zinc, inhibit biotinidase.[11] Therefore, EDTA is in-cluded in storage buffers at 1–10 mM to prevent this inhibition.

Biotinidase from hog serum and from hog liver is stable both during the enzyme assay and during preincubation for up to 30 min at 37°.[11] Denaturation of the enzyme occurs after 30 min at 60°.[11] The human serum enzyme is denatured 60% after 15 min at 60° and 100% after 15 min at 70°.[17]

17 K. Hayakawa and J. Oizumi, Clin. Chim. Acta 168, 109 (1987).

[12] Monoclonal Antibody to Biotin

By KRISHNAMURTI DAKSHINAMURTI and EDWARD S. RECTOR

Introduction

The high affinity between the glycoprotein avidin and the vitamin biotin has been exploited in various areas of biological research, particu-larly in affinity cytochemistry.[1,2] The restrictions of this method have also been recognized.[1] Various cell lines, such as HeLa, baby hamster kidney, and human skin fibroblasts, have been shown to bind and internalize avidin.[3,4] In view of this we have explored the possibility of producing monoclonal antibodies to biotin for use in studies on the cellular functions

1 E. A. Bayer and M. Wilchek, Methods Biochem. Anal. 26, 1 (1980).
2 J. Korpela, J. Med. Biol. 62, 5 (1984).
3 K. Dakshinamurti and L. E. Chalifour, J. Cell. Physiol. 107, 427 (1981).
4 L. E. Chalifour and K. Dakshinamurti, Biochem. Biophys. Res. Commun. 104, 1047 (1982).

of biotin. Antibodies produced by traditional methods are heterogeneous and hence are not true analytical reagents. We report here the production and characterization of monoclonal antibodies to biotin.

Methods

Materials. Eagle's minimal essential medium (EMEM), fetal bovine serum, penicillin G, streptomycin, fungizone, hypoxanthin, and Freund's complete adjuvant were purchased from Grand Island Biological Company (St. Louis, MO). L-[4,5-^{3}H]Leucine (62 Ci/mmol) was purchased from Amersham Corp. (Oakville, ON). Keyhole limpet hemocyanin (KLH), biotin, carbodiimide, acetyl-CoA, N-hydroxysuccinimidobiotin, ATP, bovine serum albumin (BSA), and propionyl-CoA were obtained from Sigma Chemical Co. (St. Louis, MO). [^{3}H]Biotin (10 Ci/mmol) was a gift from Hoffman-LaRoche Inc. (Nutley, NJ). Affi-Gel protein A MAPS kit and reagents for polyacrylamide gel electrophoresis were obtained from Bio-Rad Laboratories (Mississauga, ON). NaH^{14}CO$_3$ (40–60 mCi/mmol) and Aquasol-2 scintillant were purchased from New England Nuclear Corp. (Montreal, QE). Horseradish peroxidase-conjugated affinity-purified goat antimouse IgG + IgM (H + L) was obtained from Kirkeguard & Perry Labs (Gaithersburg, MD). Pristane (1,6,10,14-tetramethylpentadecane) was purchased from Aldrich Chemical Co. (Milwaukee, WI). All other chemicals used were of reagent grade.

Antigen Preparation. Biotin is covalently attached to KLH. Activated biotin (N-hydroxysuccinimidobiotin, 10 mg) is added to 62 mg of KLH dissolved in 5.0 ml of 0.2 M borate buffer (pH 9.4). This is stirred for 6 hr at room temperature and dialyzed for 24 hr against four changes of 4 liters each of 0.9% (w/v) NaCl. The biotin–KLH is used as the antigen.

Biotin is covalently attached to BSA according to the procedure of Berger.[5] Biotin (500 mg) and 5 μCi of [^{3}H]biotin are suspended in 7.5 ml of 50% aqueous pyridine with constant stirring at room temperature. Carbodiimide (2.5 g) is dissolved in 12.5 ml of 50% aqueous pyridine, and this mixture is added dropwise to the biotin suspension. Stirring is continued for an additional 30 min at room temperature during which time a clear solution is formed. Albumin (250 mg), dissolved in 6.25 ml of water, is added dropwise to the biotin–carbodiimide reaction mixture and stirring is continued at room temperature for 4.5 hr. The mixture is dialyzed for 24 hr against four changes of 4 liters each of 0.9% NaCl. Determination of the radioactivity before and after dialysis indicated approximately 8 mol of biotin bound per mole of BSA.

[5] M. Berger, this series, Vol. 62, p. 319.

Immunization. Female BALB/c mice are immunized by two subcutaneous injections of 5 μg of biotin–KLH emulsified in Freund's complete adjuvant (0.1 ml of emulsion per injection per mouse) at a 1-month interval. At 3 weeks after the second injection and 3 days before fusion, the mice are injected intraperitoneally with 5 μg of biotin–KLH in 0.5 ml of phosphate-buffered saline (PBS). At the time of killing, serum is collected to provide a positive control in subsequent immunoassays.

Cell Fusion and Cloning Procedures. The cell fusion procedure using NS-I myeloma cells and polyethylene glycol as the fusing agent has been described in detail elsewhere.[6] Six individual fusions were performed, generating a total of 576 primary cultures. Supernatants are tested for the mouse antibiotin antibodies 11 days following fusion. Selected cultures are subsequently expanded and retested for antibody activity. Of 150 cultures we produced, 4 were selected for the derivation of monoclonal cell lines. Using the limiting dilution procedure, a total of 188 clones were obtained subsequently, of which 20 were selected for expansion and storage in liquid N$_2$. Four of these cell lines, each derived originally from a different fusion, were chosen for the production of the monoclonal antibodies reported in this chapter.

Enzyme-Linked Immunosorbent Assay (ELISA). The initial screening and selection of culture supernatants containing antibodies to biotin are accomplished by ELISA. Biotin–BSA (0.1 ml), at a concentration of 10 μg/ml in Na$_2$CO$_3$ buffer (pH 9.0), is used to coat the wells of 96-well flat-bottomed polystyrene microtiter plates (Costar, Cambridge, MA). After incubation at 4° overnight, the wells are washed once and incubated for 5 min with a solution of 1% (w/v) BSA in phosphate-buffered saline (PBS/BSA). The plates are then inverted and shaken; and immediately thereafter culture supernatants or control serum samples are placed into the wells (0.1 ml/well), and the plates are incubated at room temperature for 2 hr. Following removal of the samples and one wash with 20 mM imidazole-buffered saline (IBS) containing 0.02% (v/v) Tween 20 (IBS/Tween), peroxidase-conjugated affinity-purified goat antimouse IgG + IgM (H + L) diluted 1:100 in PBS/BSA is placed into the wells (0.1 ml/well). The plates are then incubated at room temperature for 2 hr with gentle shaking. The wells are washed 5 times with IBS/Tween, and substrate solution containing 2,2'-azinodi(3-ethylbenzthiazoline sulfonate) and H$_2$O$_2$ in cacodylate buffer is added (0.1 ml/well). The apparent absorbance at 405 nm generated in each well is determined after a final incubation for 10 min at room temperature. Antiserum obtained from

[6] E. Rector, T. Nakajima, C. Rocha, D. Duncan, D. Lestourgeon, R. S. Mitchell, J. Fischer, A. H. Sehon, and G. Delespesse, *Immunology* 55, 481 (1985).

immunized mice provides a positive control for all assays. Normal mouse serum and tissue culture medium are used as negative controls.

Saturation Analysis by ELISA. Monoclonal antibody-containing supernatants from hybridoma cultures are mixed with 100 μl of varying concentrations of biotin–BSA (0–1000 $\mu g/ml$), biotin–KLH (0–1000 $\mu g/ml$), biotin (10^{-2}–10^{-10} M), and biocytin (10^{-2}–10^{-10} M). After incubation at 4° for 24 hr, the antibody–antigen mixtures are transferred into 96-well microtiter plates that had been previously coated with biotin–BSA. After incubation for 24 hr at 4°, unbound proteins are removed by three 2-min washes with IBS/Tween. The second antibody is added and the ELISA completed as described above.

Production and Purification of Monoclonal Antibodies. Female BALB/c mice are primed by intraperitoneal injection (0.5 ml/mouse) of pristane. After 2 weeks, monoclonal hybridoma cells (5×10^6) are injected intraperitoneally in 0.5 ml of culture medium. Ascites fluids are collected after approximately 10 days, placed in a tube containing heparin, and clarified by centrifugation.

Immunoglobulins from the ascites fluid are purified using the Affi-Gel protein A MAPS system according to the manufacturer's instructions. The Affi-Gel protein A column (0.5 × 10 cm) is equilibrated with 5 bed volumes of binding buffer. Ascites fluid (2 ml) is diluted with an equal volume of binding buffer and applied to the column. The column is washed with 10 volumes of binding buffer. The IgG is eluted by washing the column with 15 volumes of elution buffer. Protein-containing fractions are pooled and dialyzed against 10 mM Tris-HCl buffer (pH 7.5). Purity of the antibody is estimated by electrophoresis on 10% (w/v) sodium dodecyl sulfate (SDS)-polyacrylamide gels according to the method of Laemmli.[7]

Inhibition of Propionyl-CoA Carboxylase by Monoclonal Antibodies to Biotin. The ability of monoclonal antibodies raised against biotin to inhibit the activity of the biotin-containing enzyme propionyl-CoA carboxylase (EC 6.4.1.3) is investigated by using a 30% $(NH_4)_2SO_4$ fraction of rat liver cytosol.[8] A sample of rat liver cytosol is incubated for 15 min at 37° with ascites fluid containing monoclonal antibody to biotin. Propionyl-CoA carboxylase activity is then measured according to procedure of Lane and Halenz.[9]

One unit of enzyme activity is defined as that amount which catalyzes

[7] U. K. Laemmli, *Nature (London)* 227, 680 (1970).
[8] P. M. Gillevet and K. Dakshinamurti, *Biosci. Rep.* 2, 841 (1982).
[9] M. D. Lane and D. R. Halenz, this series, Vol. 5, p. 576.

the carboxylation of 1 μmol of substrate per minute at 37° into acid-stable product.

Protein Measurement. Protein content is determined according to the procedure of Lowry et al.[10] Bovine serum albumin is used to construct the standard curve.

Development of Monoclonal Antibodies

Fusion of spleen cells obtained from mice immunized with biotin–KLH with the murine myeloma cell line NS-1 have yielded a highly successful fusion. The supernatants from initial fusion cultures were all strongly positive for antibiotin antibodies, and of the crude cultures chosen for secondary screening all retained the ability to produce antibiotin antibodies. Limiting dilutions of crude cultures yielded approximately 50% of the possible number of single clones (188/384) and of these 75% were positive for antibiotin antibodies. Monoclonal antibody was obtained by ascites production in mice using four different clones (Nos. 28, 33, 38, and 142). One clone, No. 33, produces antibody of high affinity that binds both free and haptenic biotin antigens (Fig. 1A–C) as well as biocytin (Fig. 1D). Figure 1 shows that both biotin–BSA and biotin–KLH inhibit antibody binding to the solid phase at concentrations several orders of magnitude lower than that of free biotin.

The profiles of the binding of biotin–BSA and biotin–KLH are similar, and on that basis they should have an equal number of binding sites. It is not possible to calculate the molar ratios of binding between free biotin and biotin–KLH as can be done for free biotin and biotin–BSA, in view of the wide range (3×10^5–9×10^6) given for the molecular mass of hemocyanin. However, the fact that protein-bound biotin inhibits antibody binding to the solid phase more effectively, on a molar basis, than does free biotin should facilitate biological use of this antibody, since cellular and circulating biotin is mostly protein bound. Antibody produced by clone No. 33 is also highly effective at inhibiting the activity of the biotin-containing enzyme propionyl-CoA carboxylase (Table I). The other clones produce antibody of much lower affinity than that derived from clone 33 and are 1000-fold less effective at inhibiting propionyl-CoA carboxylase activity. The activity of acetyl-CoA carboxylase is inhibited in a fashion similar to that for propionyl-CoA carboxylase [4.25×10^3 milliunits (mU) of acetyl-CoA carboxylase is inhibited by 1 ml of ascites fluid

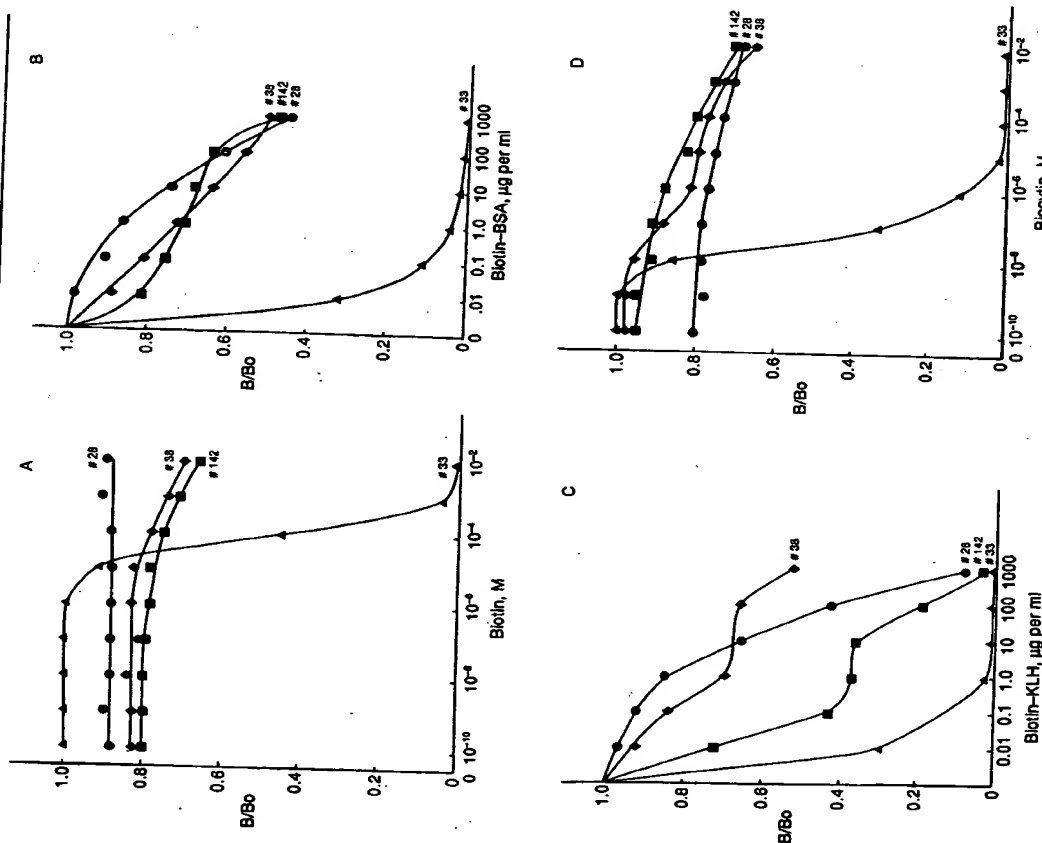[10] O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.* 193, 265 (1951).

TABLE I

INHIBITION OF PROPIONYL-CoA CARBOXYLASE
BY MONOCLONAL ANTIBODIES DIRECTED
TOWARD BIOTIN[a]

Clone no.	Propionyl-CoA carboxylase (mU) inhibited by 1 ml of ascites fluid
28	14.6
33	15.2×10^3
38	10.4
142	14.9

[a] Rat liver propionyl-CoA carboxylase was incubated for 15 min at 37° with ascites fluid containing monoclonal antibody to biotin. At the end of the incubation, enzyme activity was measured as described in the text. From K. Dakshinamurti, R. P. Bhullar, A. Scott, E. S. Rector, G. Delespesse, and A. Sehon, *Biochem. J.* **237**, 477 (1986), with permission of the publishers.

from clone No. 33]. There is no inhibition of either carboxylase when they were incubated with ascites fluid which did not contain antibiotic antibody.

Purification of Antibody from Ascites Fluid

Monoclonal antibody is purified from ascites fluid produced from clone 33 using the Affi-Gel protein A MAPS system. This system results in a one-step purification of antibody from the ascites fluid. SDS-polyacrylamide gel electrophoretic analysis showed only two major protein bands, corresponding to the immunoglobulin heavy and light chains, thus confirming the high purity of the final antibody preparation. The purified antibody was effective at inhibiting both acetyl-CoA carboxylase and propionyl-CoA carboxylase activities [2.59 mU of acetyl-CoA carboxylase and 2.33 mU of propionyl-CoA carboxylase, respectively, inhibited per microliter of antibody (0.13 μg/protein)].

Comments

Immune serum containing antibodies to the haptenic group biotin has previously been prepared by immunization of rabbits with biotin–BSA.[11]

[11] F. Ahmad, P. M. Ahmad, R. Dickstein, and E. Greenfield, *Biochem. J.* **197**, 95 (1981).

FIG. 1. Relative affinities of monoclonal antibiotin antibodies for biotin (A), biotin–BSA (B), biotin–KLH (C), and biocytin (D). The relative affinity of monoclonal antibodies (clone numbers indicated) for different biotin-containing antigens was determined by competition saturation analyses. The monoclonal antibodies were incubated overnight at 4° with varying concentrations of antigen prior to performing an ELISA with biotin–BSA as a plate-coating antigen. B/B_0 represents the amount of monoclonal antibody bound to the plates in the presence of competing antigen (B) relative to the amount bound in the absence of competing antigen (B_0). [From K. Dakshinamurti, R. P. Bhullar, A. Scoot, E. S. Rector, G. Delespesse, and A. Sehon, *Biochem. J.* **237**, 477 (1986), with permission of the publishers.]

Biotin transport as well as the nonprosthetic group function of biotin could be studied by using the specific antibody. Antibiotin antibody could also be used in isolation and purification procedures where the affinity between biotin and avidin proves to be too strong to be useful.

Acknowledgment

However, the multivalent specificity of these antibodies makes them a less attractive reagent in assessing the role of biotin in metabolism. This led us to develop a monoclonal antibody to biotin.

The monoclonal antibody can be used in assessing the role of the biotinyl group in biotin-containing enzymes. Acetyl-CoA carboxylase of mammalian origin exists in the cell as an inactive protomer, which is converted to the enzymically active form in the presence of citrate.[12] As the monoclonal antibody to biotin can be taken up by HeLa cells in culture, the antibody can be used to investigate the cellular protomer-to-polymer ratio of acetyl-CoA carboxylase under different nutritional and hormonal conditions. Assay in vitro for acetyl-CoA carboxylase has shown that the polymerized enzyme is not inactivated by the monoclonal antibody whereas it is inactivated when treated with antibody prior to polymerization.

The high affinity between biotin and avidin has been utilized in various areas of biological research. Of much significance is the use of the avidin–biotin complex for the localization and evaluation of cell surface receptors. In these studies, biotin is attached to cell surface functional groups through an appropriate reactive derivative. It has been assumed that coupling a small molecule such as biotin to the receptor would only slightly affect binding characteristics. In some instances this has been verified prior to the use of this technique. However, this is not always so. Thus, the addition of avidin to the medium of biotinylated lymphocytes has been used to mimic lectin-induced stimulation. Whereas the lectin-induced stimulation of lymphocytes is reversible, the stimulation observed with avidin was irreversible.[1]

There are other possible limitations in using this system. Before using the versatile avidin–biotin complex method, it must be determined whether the interaction in the experimental system under study comprises a biotin-containing, biotin-recognizing, or biotin-free system. Moreover, avidin is a basic glycoprotein[13]; the oligosaccharide moiety or its ion-exchange properties might thus contribute to side reactions. In this regard, it has been shown that avidin, perhaps mimicking a natural ligand, binds to isolated rat liver plasma membranes[14] and to a variety of cells studied.[3,4] This emphasizes the need for proper controls for applying the avidin–biotin complex technique to affinity cytochemical studies. The use of monoclonal antibody in these studies would obviate such problems.

[12] M. D. Lane, J. Moss, and S. E. Polakis, Curr. Top. Cell. Regul. 8, 139 (1974).
[13] N. M. Green, Adv. Protein Chem. 29, 85 (1975).
[14] L. E. Chalifour and K. Dakshinamurti, Biochem. J. 210, 121 (1983).

Developmental patterns of free and protein-bound biotin during maturation and germination of seeds of *Pisum sativum*: characterization of a novel seed-specific biotinylated protein

Manuel DUVAL, Claudette JOB, Claude ALBAN, Roland DOUCE and Dominique JOB*

Unité Mixte C.N.R.S./Rhône-Poulenc (Unité associée au Centre National de la Recherche Scientifique, U.M. 41), Rhône-Poulenc Agrochimie, 14-20 rue Pierre Baizet, 69263 Lyon Cédex 9, France

Mature dry pea seeds contain three major biotinylated proteins. Two of these of subunit molecular mass about 75 kDa and 200 kDa are associated with 3-methylcrotonyl-CoA carboxylase (EC 6.4.1.4) and acetyl-CoA carboxylase activities (EC 6.4.1.2) respectively. The third does not exhibit any of the biotin-dependent carboxylase activities found in higher organisms and represents the major part of the total protein-bound biotin in the seeds. This novel protein has been purified from a whole pea seed extract. Because in SDS/polyacrylamide gels the protein migrates with an apparent molecular mass of about 65 kDa, it is referred to as SBP65, for 65 kDa seed biotinylated protein. The molecular mass of native SBP65 is greater than 400 kDa, suggesting that the native protein assumes a polymeric structure, resulting from the association of six to eight identical subunits. The results of CNBr cleavage experiments suggest that biotin is covalently bound to the protein. The stoichiometry is 1 mol of biotin per 1 mol of 65 kDa polypeptide. The temporal and spatial pattern of expression of SBP65 is described. SBP65 is specifically expressed in the seeds, being absent from leaf, root, stem, pod and flower tissues of pea plants. The level of SBP65 increases dramatically during seed development. The protein is not detectable in very young seeds. Its accumulation pattern parallels that for storage proteins, being maximally expressed in the mature dry seeds. SBP65 disappears at a very high rate during seed germination. The level of free biotin has also been evaluated for various organs of pea plants. In all proliferating tissues examined (young developing seeds, leaf, root, stem, pod and flower tissues), free biotin is in excess of protein-bound biotin. Only in the mature dry seeds is protein-bound biotin (i.e. that bound to SBP65) in excess of free biotin. These temporal expression patterns, and the strict organ specificity for expression of SBP65, are discussed with regard to the possibility that in plants, as in mammals, biotin plays a specialized role in cell growth and differentiation.

INTRODUCTION

In all organisms biotin is an essential cofactor for a small number of enzymes involved in CO_2 transfer during carboxylation reactions [1]. Although the existence of this vitamin is well recognized in plants, little is known about its biosynthesis and function. It is known that animal cells contain four biotinylated enzymes, acetyl-CoA carboxylase (ACC; EC 6.4.1.2), 3-methyl-crotonyl-CoA carboxylase (MCC; EC 6.4.1.4), propionyl-CoA carboxylase (PCC; EC 6.4.1.3) and pyruvate carboxylase (PC; EC 6.4.1.1). They play central roles in a variety of metabolic and catabolic processes, supporting essential cellular housekeeping functions. ACC, which catalyses the ATP-dependent carboxylation of acetyl-CoA, is recognized as the regulatory enzyme of lipogenesis. MCC catalyses the conversion of 3-methylcrotonyl-CoA into 3-methylglutaconyl-CoA, a key reaction in the degradation pathway of leucine. PCC is a key enzyme in the catabolic pathway of odd-chain fatty acids, isoleucine, threonine, methionine and valine. PC has an anaplerotic role in the formation of oxaloacetate [2].

It is now established that plant cells contain the four bio-tinylated enzymes found in animals [3]. The most extensively studied biotin carboxylase is ACC, because of its obvious function in membrane biogenesis, and because the enzyme isolated from vegetative tissues and developing seeds of monocotyledonous plants is the target of the very potent cyclohexanedione and aryloxyphenoxypropionate herbicides [4–7]. There is evidence suggesting that biotin and biotin-containing proteins might play specialized roles in regulation of plant development. Thus ACC is required for both growth of vegetative tissues and synthesis of storage lipids in developing seeds [8,9]. MCC was found to increase rapidly during pea leaf development [10]. Also, during development of carrot somatic embryos, a 50-fold increase in the level of a 62 kDa biotinylated polypeptide has been observed, as embryogenic cell clusters developed into torpedo embryos [11]. Attention has been focused recently on the existence of free biotin in plant cells. A study conducted with pea leaves showed the existence of a free biotin pool in the cytosolic compartment, accounting for about 90 % of the total (free plus protein-bound) biotin. It has been argued that this pool might control the expression of genes encoding biotinylated carboxylases and/or enzymes involved in biotin synthesis [12]. The role of biotin in plants is also best illustrated by the discovery of a mutation that causes defective embryo development in *Arabidopsis thaliana* and requires the vitamin at a critical stage of embryogenesis. Thus no biotin was detectable in the arrested embryos of the mutant, and mutant embryos were specifically rescued when grown in the presence of biotin [13,14].

As a means of understanding the function of this vitamin in developing and germinating seeds, we have analysed the content

---

of free and protein-bound biotin of pea seeds at various developmental stages. A detailed comparative study was also carried out with various organs of pea plants (leaves, roots, stems, flowers and pods). The general finding is that in mature dry seeds the level of bound biotin is higher than that of free biotin, owing to the existence of a seed-specific biotinylated 65 kDa polypeptide which we called SBP65. This protein does not support *in vitro* any of the four biotin-dependent carboxylase activities present in higher organisms. This contrasts with the situation observed in all developing organs, including young developing seeds and pea plants, where the free vitamin is present in higher amounts than that bound to proteins. The possible implications of these findings are discussed.

MATERIALS AND METHODS

Reagents

ATP, acetyl-CoA, 3-methylcrotonyl-CoA, D-biotin, biotin-labelled β-galactosidase, biotin-labelled molecular-mass markers, anti-biotin peroxidase-conjugated antibodies, horseradish peroxidase-conjugated streptavidin, 2,2'-azinobis(3-ethylbenzo-thiazoline-6-sulphonic acid) and 4-chloro-1-naphthol were from Sigma. $NaH^{14}CO_3$ (53.1 mCi/mmol) and D-[*carbonyl*-^{14}C]biotin (56.6 mCi/mmol) were from Amersham. All other chemicals were of analytical grade. Solutions were made in ultrapure water (Millipore) and filtered with 0.2 μm filters.

Plant material and samples for developmental studies

Mature dry pea seeds (*Pisum sativum*, cv. Douce Provence) used in this study are referred to as S0 seeds. They had a mean fresh weight of 250 mg and a diameter of about 6–7 mm. Plants were grown from these seeds in soil under a 12 h photoperiod of white light from fluorescent tubes (10–40 μE·ml^{-2}·s^{-1}) at 20 °C. They were watered every day with tap water. Different organs (leaves, roots, stems, flowers, pods and seeds) were harvested at different times, and stored at -75 °C until use. No germination was visible at day 1 after planting. The mean fresh weight of the 1-day-old seeds was 475 mg. Germination occurred between 36 h (0 % of seed germination) and 48 h (approx. 90 % of seed germination) after sowing. The mean fresh weight of the 2-day-old seeds was 528 mg. At day 2 and day 3, mean fresh weights of the seedlings (minus cotyledons) were 62 and 120 mg respectively. Flowering started at day 25. During the initial period of seed formation seed samples showed heterogeneity in size. S36(1), S36(2) and S36(3) refer to seeds collected from pods at day 36. Mean fresh weights per seed were 7, 32 and 144 mg respectively. Samples sizes were approx. 1.5, 2.5 and 3.5 mm diameter, respectively. S41, S52 and S76 refer to seeds collected at day 41, 52 and 76. Mean fresh weights per seed were 300, 500 and 210 mg respectively. The size of these three samples was approx. 6–7 mm diameter. At day 76, seeds were fully matured and desiccated.

Preparation of extracts for development studies

Plant materials (1–5 g) were frozen in liquid nigrogen and finely ground using a mortar and pestle. The powder was homogenized in 2 vol of buffer A [50 mM Hepes, pH 8.0, 10 % (v/v) glycerol, 1 mM EDTA, 5 mM dithiothreitol (DTT), 1 mM phenyl-methanesulphonyl fluoride, 1 mM benzamidine/HCl, 5 mM 6-aminohexanoic acid], followed by centrifugation (40000 g,

30 min). The supernatant comprised the crude extract which was processed as follows. (i) Portions of the extract were used for the determination of total (free plus protein-bound) biotin. (ii) To 1 ml of crude extract, 4 ml of cold acetone (-20 °C) was added, and the mixture left to stand for 30 min to precipitate proteins. After centrifugation (3000 g, 10 min), the supernatant was evaporated to dryness under a stream of nitrogen, and 1 ml of PBS containing 0.1 % (v/v) Tween 20 (PBST) was added. This solution was used for quantification of free biotin. Controls, in which D-[^{14}C]biotin was added to the extracts before the addition of acetone, established that the radioactivity was quantitatively recovered from the supernatant fraction. Also, it was verified that the supernatants were depleted of protein-bound biotin. (iii) Crystalline $(NH_4)_2SO_4$ was added to each crude extract with stirring until 50 % (w/v) saturation was achieved. The mixture was stirred for 30 min at 4 °C, and then centrifuged (40000 g, 20 min). The pellet was resuspended in 2 vol. of buffer A, and the solution was used for biotin carboxylase activity measurements and for quantification of protein-bound biotin. For SDS/PAGE analyses, 1 vol. of SDS sample buffer [10 mM Tris/HCl, pH 6.8, 2.5 % (w/v) SDS, 15 % (v/v) 2-mercaptoethanol, 30 % glycerol, 0.06 % (w/v) Bromophenol Blue] was added to 2 vol. of each sample and heated to 100 °C for 5 min.

Protein determination, activity measurements and electrophoresis

Protein was quantified by the method of Bradford [15] with BSA as the standard. Biotin carboxylase activities were measured as the incorporation of radioactivity from $NaH^{14}CO_3$ into an acid-stable product [10,11]. Assays contained 50 mM Hepes, pH 8.0, 2.5 mM $MgCl_2$, 1 mM ATP, 2 mM DTT, 10 mM $NaH^{14}CO_3$ (1 mCi/mmol), 20 mM KCl, 0.4 mM appropriate substrate (acetyl-CoA, 3-methylcrotonyl-CoA, propionyl-CoA or pyruvate) and 1–150 μg of protein sample, in a final volume of 200 μl. Incubations were for 15 min at 30 °C. One unit of enzyme activity is equivalent to the incorporation of 1 nmol of $^{14}CO_2$ into acid-stable product in 1 min at 30 °C. SDS/PAGE was conducted with a PhastSystem (Pharmacia) in preformed gels (PhastGel Pharmacia) containing 12.5 % (w/v) acrylamide. Proteins were electrotransferred from the gel on to nitrocellulose (Bio-Rad), using a PhastTransfer device (Pharmacia), as recommended by the manufacturer. Blots were incubated for 1 h at 25 °C in PBS containing 3 % (w/v) BSA, then for 1 h with peroxidase-conjugated streptavidin (250 ng/ml) in PBST, followed by three washes with PBST. Biotinylated polypeptides were revealed by adding peroxidase/substrate solution (1 μl of 30 % H_2O_2/ml and 0.5 mg of 4-chloro-1-naphthol/ml).

Streptavidin-binding assays for biotin

Since streptavidin binds both free and protein-bound biotin [16], two types of assays were used.

Direct solid-phase biotin assay for protein-bound biotin

Microtitre e.l.i.s.a. plates (Greiner) were incubated for 3 h at 25 °C with various amounts of sample (0.01–10 μg of protein per well) in 100 μl of PBS, washed four times with PBST, and each well received 50 ng of peroxidase-conjugated streptavidin in 100 μl of PBST. After incubation for 1 h at 25 °C, wells were washed four times with PBST, and each received a 100 μl peroxidase/substrate solution [0.1 μl of 3 % H_2O_2 and 100 μg of 2,2'-azino[bis-(3-ethylbenzothiazoline-6-sulphonic acid)] in 100 mM citrate phosphate, pH 4.0. Colour development was quantified with an EL-312 microplate reader (Biotek Instruments),

using a 405 nm filter. Standard curves were generated from control wells coated with known amounts (1–20 ng) of biotin-labelled β-galactosidase.

Indirect solid-phase biotin assay for free and protein-bound biotin

Control experiments showed that free D-biotin cannot bind to the wells of the microtitre plates. Free biotin was quantified by a modification of standard procedures [13,17]. E.l.i.s.a. plates were incubated for 3 h at 25 °C with a fixed amount of biotin-labelled β-galactosidase (usually 100 ng per well) in 100 μl of PBS, and then washed four times with PBST. To generate a standard curve, serial dilutions (0.002–80 ng) of D-biotin (final volume 100 μl in PBST) were prepared in Eppendorf centrifuge tubes. They were mixed with 100 μl of PBST containing a fixed amount (usually 20 ng) of peroxidase-conjugated streptavidin, followed by incubation for 1 h at 25 °C. This yielded a series of mixtures containing various amounts of free and biotin-bound peroxidase-conjugated streptavidin. Portions (100 μl) of these mixtures were then transferred to each biotinylated β-galactosidase-coated well of the plates. After incubation for 1 h at 25 °C, plates were processed as for the direct solid-phase biotin assay, i.e. they were washed four times with PBST and, after addition of the peroxidase/substrate solution, A_{405} was measured as above. Because only free peroxidase-conjugated streptavidin will bind to biotinylated β-galactosidase, and because exchange of biotin from preformed biotin–streptavidin complexes is extremely slow [16], this yielded a titration curve, A_{405} versus [D-biotin], such that (i) the lower the [biotin] the higher the A_{405} and (ii) at high [biotin] the A_{405} was close to the background values. The point at which the slope undergoes a sharp increase indicates the equivalent point, i.e. the minimal amount of biotin required to saturate the fixed amount of conjugated streptavidin used in the assay. The concentration of free D-biotin at the equivalence point was not dependent on the duration of the incubation with peroxidase substrates, although prolonged incubations increased the sensitivity. Furthermore, the concentration of free D-biotin at the equivalence point was dependent on the concentration of conjugated streptavidin, but not on that of biotin-labelled β-galactosidase in the assays. The sensitivity was of the order of 2 pg of biotin.

To quantify free biotin in the various plant extracts, the same procotol was used except that the serial dilutions of free D-biotin used to construct the calibration curve were replaced by serial dilutions in PBST of the supernatants derived from the acetone-treated plant extracts. Free biotin content in these samples was then calculated from the equivalence points, with reference to the calibration curve, obtained under the same conditions. In principle, protein-bound biotin could also be quantified by this assay from the precipitate of the acetone-treated samples. Yet, attempts to solubilize the acetone pellet from the various extracts under conditions compatible with preservation of streptavidin-binding activity (e.g. in the absence of denaturing agents such as SDS) produced unreliable results. The same limitations were encountered when proteins in the crude extracts were precipitated with trichloroacetic acid. The unprocessed crude extracts and the corresponding acetone supernatants were used therefore for total (free plus protein-bound) and free biotin determinations respectively. Whenever possible, the level of protein-bound biotin was then calculated from the difference between these two quantities.

Purification of SBP65 from mature pea seeds

All purification steps were carried out at 4 °C. Chromatographic steps were performed with a Pharmacia f.p.l.c. system. Frozen (−75 °C) mature pea seeds (100 g) were finely ground in a Waring blender. To the powder 500 ml of buffer A was added, and the mixture was homogenized with a Polytron homogenizer. After centrifugation (15000 g, 30 min), the resulting supernatant (17 g of protein) was brought to 200 g/l $(NH_4)_2SO_4$, stirred for 30 min at 4 °C, allowed to stand at this temperature for 2 h, and then centrifuged (15000 g, 30 min). The pellet (4 g of protein) was resuspended in 200 ml of buffer A. This protein extract was brought to 50 g/l poly(ethylene glycol) (PEG) 6000, stirred for 20 min at 4 °C, and centrifuged (15000 g, 30 min). The derived supernatant was brought to 200 g/l PEG 6000, stirred for 20 min at 4 °C, and centrifuged as above. The sticky brown-coloured pellet was resuspended in 15 ml of buffer A with an Ultra-Turrax homogenizer; then the suspension was clarified by centrifugation (50000 g, 30 min), and the supernatant (0.75 g protein, 12 ml) filtered on a 0.2 μm filter. At this step, biotinylated proteins had been purified about 22-fold with a 60 % recovery. This extract was subjected to monomeric-avidin affinity chromatography [10, 18–20]. Sample was loaded (flow rate 0.1 ml/min) on the column (1.5 cm × 20 cm) equilibrated in buffer B (20 mM Hepes, pH 8.0, 10 % glycerol, 0.5 M KCl, 1 mM EDTA, 1 mM benzamidine/HCl, 5 mM 6-aminohexanoic acid). After the column had been washed with 350 ml of buffer B (flow rate 0.2 ml/min), bound proteins (0.8 mg) were eluted with 50 ml of buffer B containing 2 mM D-biotin. The yield for protein-bound biotin was in the range 20–30 %. Eluted fractions were dialysed against buffer B containing 50 mM KCl, concentrated to 6 ml with Macrosep-10 tubes (Filtron) and applied to a Mono Q HR5/5 column (0.5 cm × 5 cm; Pharmacia) equilibrated in buffer B containing 50 mM KCl. SBP65 was recovered in the non-absorbed fraction. After the column had been washed with 20 ml of buffer B containing 0.1 M KCl, adsorbed proteins, supporting ACC and MCC activities, were eluted with 10 ml of buffer B containing 0.3 M KCl (flow rate 0.5 ml/min; fraction size 0.5 ml). Molecular-mass determinations were effected by gel filtration on a Sephacryl S-300 HR column (2.6 cm × 35 cm, 180 ml; Pharmacia). Protein samples were allowed to react with CNBr as described [21].

Preparation of antiserum

Purified SBP65 (800 μg) was subjected to SDS/PAGE. After electroelution from the gel with 25 mM Tris/192 mM glycine, pH 9.0, containing 0.1 % SDS, the protein was dialysed against 50 mM Tris/HCl, pH 7.8. Antibodies directed against the electroeluted SBP65 (anti-SBP65) were raised in a rabbit by standard protocols. For the preparation of affinity-purified antibodies, SBP65 (30 μg) was subjected to SDS/PAGE and electrotransferred on to nitrocellulose. A thin strip corresponding to SBP65 was excised. After being washed with PBS containing 3 % BSA, the strip was incubated with a 20-fold dilution of the anti-SBP65 serum, overnight at 4 °C. After three washes with PBST, specific antibodies were eluted by a quick (30 s) wash of the strip with 1 ml of 0.2 M glycine/HCl, pH 2.2, and then immediately neutralized by adding 170 μl of 1 M Tris/HCl, pH 8.8 [22].

D-[^{14}C]Biotin binding and exchange

Assays contained 50 mM Hepes, pH 8.0, 1 mM DTT, 1.6 μM D-[^{14}C]biotin (56.6 mCi/mmol) and sample (10–200 μg of protein), in a total volume of 15 μl. Assays were incubated at 25 °C for 30 min, and then portions (9 μl) of reaction mixtures were spotted on glass-fibre filters. After eight washing steps in 5 % trichloroacetic acid, the filters were dried, and the trichloroacetic-
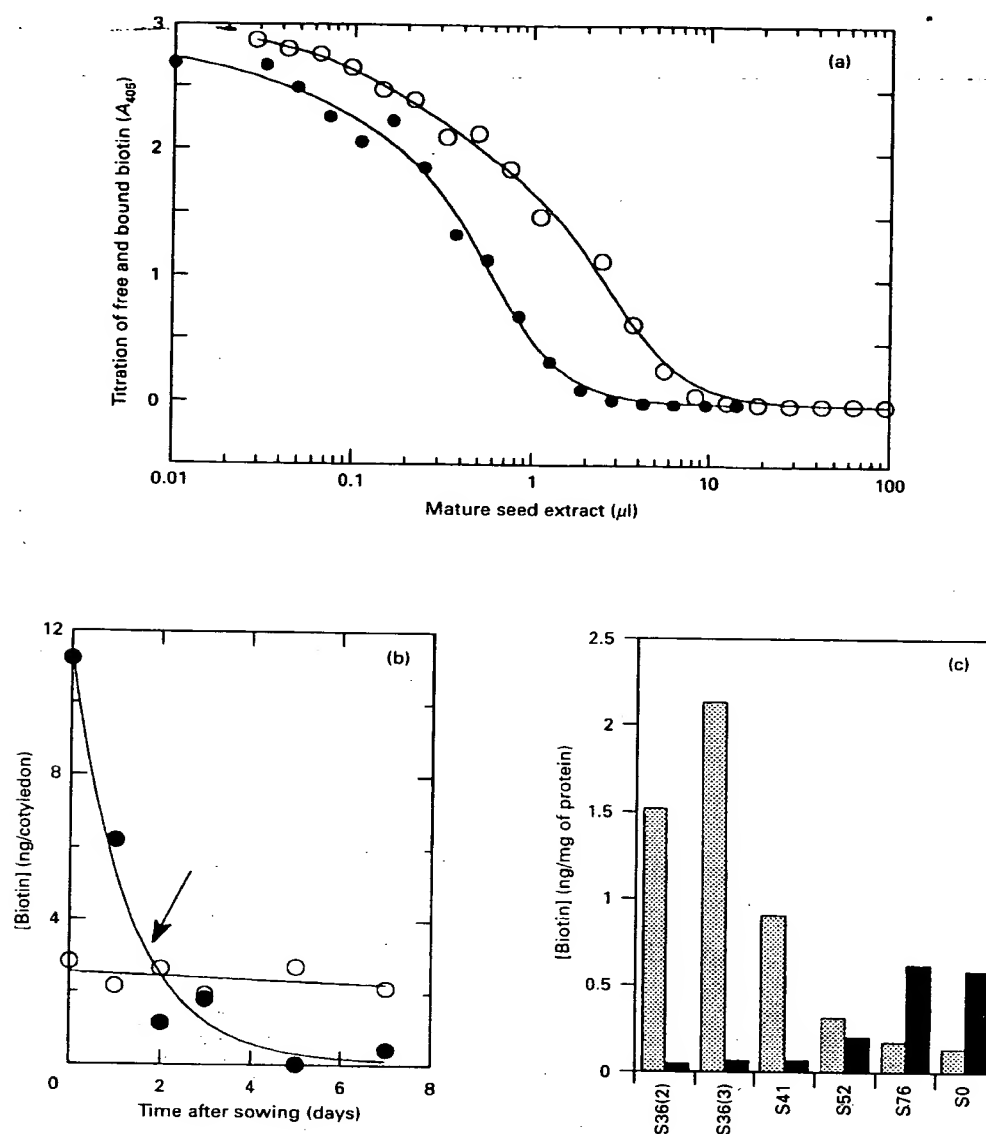
Figure 1   Biotin content of pea seeds

(a) Titration curves for biotin in mature dry pea seeds. A crude extract (1.5 seeds/ml; 72 mg of protein/ml) was prepared as described in the Materials and methods section. Serial dilutions of the unprocessed crude extract in PBST were analysed for total (free plus protein-bound) biotin (●) by the indirect solid-phase biotin assay as described in the Materials and methods section. The equivalence point was obtained at 1.7 μl of unprocessed extract. To 1 ml of this extract 4 ml of cold (− 20 °C) acetone was added. After centrifugation, the supernatant was saved, evaporated to dryness and resuspended in 1 ml of PBST. Serial dilutions in PBST of this acetone-derived extract were analysed for free biotin (○) by the indirect solid-phase biotin assay. The equivalence point was obtained at 9.4 μl of extract. For the calibration curve constructed as described in the Materials and methods section, the equivalence point was obtained at 71 pg of free o-biotin, allowing us to calculate amounts of 5 ng of free biotin and 28 ng of total biotin per S0 seed. These corresponded mostly to cotyledons; contributions from the seed coats and embryonic axes were negligible. (b) Evaluation at different times after sowing of the protein-bound (●) and free (○) biotin content of total seeds (day 0 and day 1) and cotyledons (days 2, 3, 5 and 7). Results are expressed on a per cotyledon basis. Total and free biotin were quantified as in (a). Bound biotin was calculated as the difference between the amounts of these two quantities. Radicle emergence occurred between 36 h and 48 h after sowing (arrow). (c) Evaluation of the protein-bound (■) and free (▨) biotin content during pea seed development expressed on a per mg of protein basis. For nomenclature of seed samples, see the Materials and methods section. Free biotin was quantified as in (a). Protein-bound biotin was estimated by using the direct solid-phase biotin assay.

acid-precipitable radioactivity was measured by liquid-scintillation counting.

RESULTS

Non-covalent protein-bound biotin in extracts from mature dry pea seeds

Biotin-containing proteins from plant extracts have usually been analysed on nitrocellulose filters of blotted gels, in a system using streptavidin as a specific reagent, analogous to Western blotting [10,11,23]. Presumably, only those proteins containing the covalently attached prosthetic group are detected by this method [23]. To quantify the bound biotin content of pea seeds, it was therefore of importance to investigate whether these seeds contained proteins involved in non-covalent binding of the vitamin. When an unprocessed crude extract from S0 seeds was
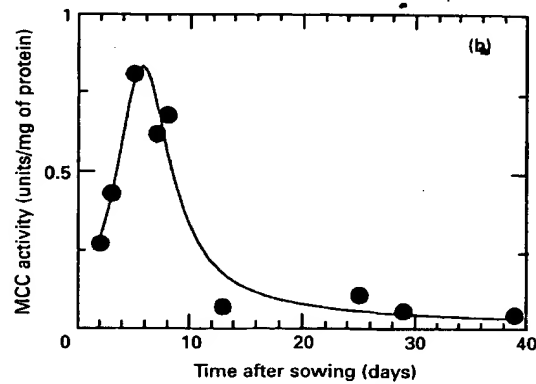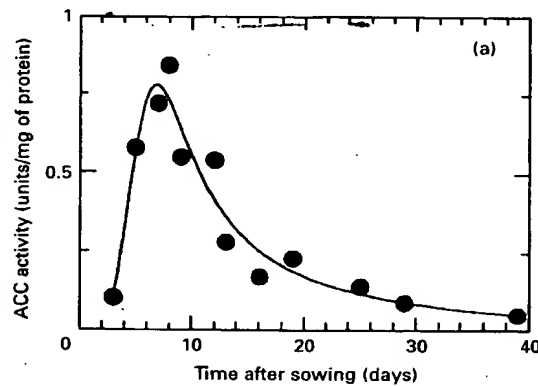
Figure 2   Requirement for biotin in early plant growth

(a) Evaluation of ACC activity in pea leaves. (b) Evaluation of MCC activity in pea roots.
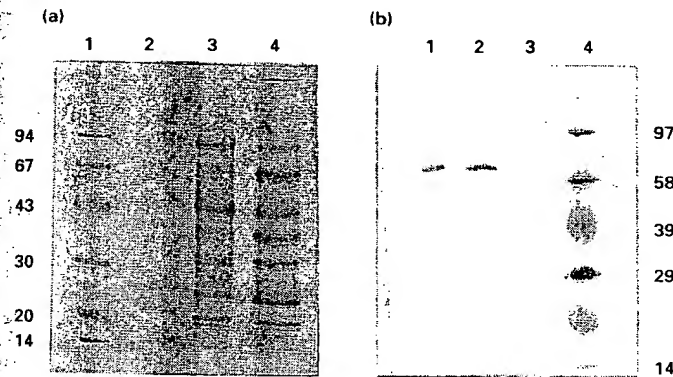

Figure 3   Distribution of SBP65 in different compartments of mature dry pea seeds

(a) Coomassie Blue R250 staining of a homogeneous SDS/polyacrylamide gel containing 12.5% acrylamide loaded with total protein extracts from seed coat (lane 2, 1 μg), embryonic axis (lane 3, 4 μg) and cotyledons (lane 4, 4 μg) of mature dry pea seeds. Molecular-mass markers (lane 1) are indicated in kDa. (b) Nitrocellulose blot of a duplicate of (a) probed with peroxidase-conjugated streptavidin. Lane 1, cotyledon extract (40 μg); lane 2, embryonic axis extract (40 μg); lane 3, seed coat extract (10 μg). Molecular-mass biotinylated markers (lane 4) are indicated in kDa.

Table 1   Purification procedure for pea SBP65

SBP65 was purified starting from 100 g of mature dry pea seeds. Protein was determined by the method of Bradford [15] with BSA as standard. Bound biotin was measured by using the direct solid-phase biotin assay described in the Materials and methods section.

Purification stage	Total protein (mg)	Total biotin (μg)
Crude extract of soluble proteins	17000	10
PEG 6000 homogenate	750	6.4
Monomeric-avidin Sepharose pool	0.8	2
Mono Q HR5/5 pool (salt-eluted fraction)	0.15	0.2
Mono Q HR5/5 pool (flow-through fraction)	0.6	2

incubated with D-[^{14}C]biotin, no radioactivity was recovered in a form precipitable by trichloroacetic acid. Under the same conditions, binding of D-[^{14}C]biotin to streptavidin could be demonstrated. Therefore, unlike egg yolk and egg white [24], mature dry pea seeds do not seem to contain significant amounts of proteins involved in non-covalent binding of biotin.

Free and protein-bound biotin in developing, mature and germinating pea seeds

Figure 1(a) shows that mature dry seeds contain a 5-fold higher amount of protein-bound biotin than free biotin. Respective amounts of about 24 ng and 5 ng of vitamin per seed were calculated, corresponding to about 100 pg of protein-bound biotin and 20 pg of free biotin per mg fresh weight. These values were within the range of total biotin levels reported for various feedstuffs of plant origin [25] and for seeds of *A. thaliana* [13].

A considerable amount of the initial protein-bound biotin disappeared during early stages of germination. At day 2, which coincides with radicle emergence, the level of bound biotin was of the order of 20% of that in the mature dry seeds. After day 5, this level was below detection (Figure 1b). Figure 1(c) shows the data obtained from seed samples collected at various development stages. On a per mg of protein basis, the level of free biotin was high in the young seeds, and then decreased during development. In contrast, the protein-bound biotin level was low in the young seeds, and then rose sharply during late stages of seed development and onset of desiccation, being maximal in the mature dry seeds.

Content of free and total biotin in various organs of pea plants

Biotin contents were determined for various organs of pea plants collected from 4 to 40 days after sowing. All extracts from leaves, roots, stems, pods and flowers yielded the same results as those obtained for the young pea seeds, i.e. although free biotin could be easily detected with the indirect solid-phase biotin assay, the amount of protein-bound biotin was almost negligible compared with that of free biotin. At day 39, free biotin levels of 11, 8.5, 7.5, 2 and 1 ng/mg of protein were calculated for the leaves, roots, stems, pods and flowers respectively. These levels varied during plant growth. For example, at day 5 and 25 the free biotin content was 1 and 9 ng/mg of protein for the leaves and 3 and 10 ng/mg of protein for the roots. This is in agreement with reported values of 1.2 ng/mg of protein and 0.2 ng/mg of protein
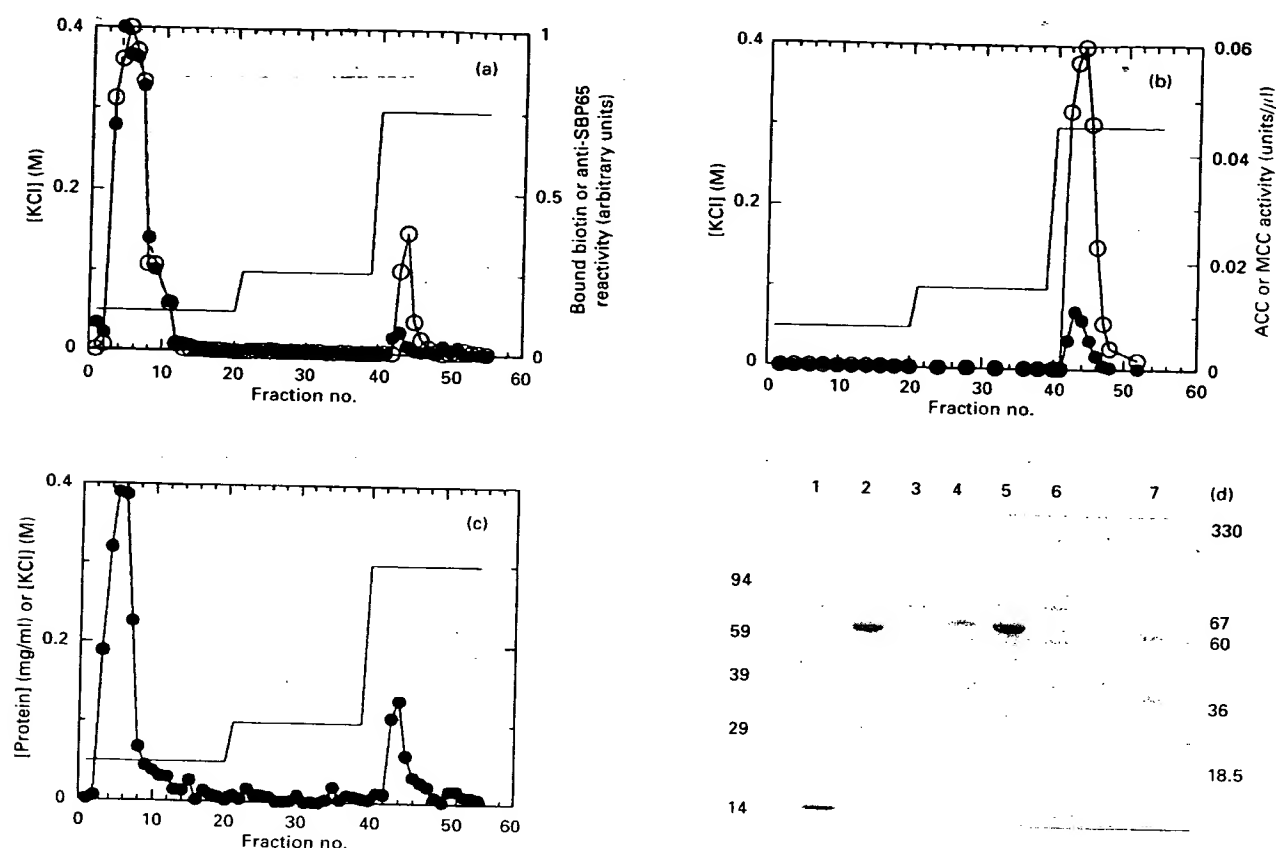
Figure 4   Purification of SBP65 by chromatography on Mono Q HR5/5

Biotinylated proteins were purified as described in the Materials and methods section, from a crude extract of mature dry pea seeds. Biotinylated fractions eluted from the monomeric-avidin affinity column were pooled (0.8 mg of protein), dialysed against buffer B containing 50 mM KCl, and loaded on a Mono Q HR5/5 column equilibrated in the same buffer. The column was washed with 20 ml of buffer B containing 0.1 M KCl, and then eluted with 10 ml of buffer B containing 0.3 M KCl. Fractions of 0.5 ml were collected, and assayed for protein-bound biotin (a, ●), affinity-purified anti-SBP65 antibody reactivity (a, ●), ACC activity (b, ●), MCC activity (b, ○) and protein (c, ●); the KCl concentration is shown (——). (d) Coomassie Blue R250 staining of an SDS/polyacrylamide gel containing 12.5% acrylamide loaded with pooled flow-through fractions 3–8 from the Mono Q column (lanes 2 and 5, 1.2 μg of protein), and pooled fractions 43–48 eluted from this column in the presence of 0.3 M KCl (lanes 3 and 6, 0.5 μg of protein). The proteins loaded on the Mono Q column are shown (lane 4, 0.8 μg of protein). Molecular-mass markers (lanes 1 and 7) are indicated in kDa.

for the free and protein-bound biotin content of 10-day-old pea leaves respectively [12].

Although the protein-bound biotin content was low, extracts from leaves, roots, stems, pods and flowers contained significant levels of ACC and MCC activities. Those for PCC and PC were close to background values. ACC activity increased strongly in the leaves from day 3 to day 8, and then declined in older tissue (Figure 2a), similarly to previous observations with maize leaves [26]. The same pattern of behaviour was observed for MCC in the roots (Figure 2b). These data suggest that there is a strong demand for biotin in the early stages of plant growth.

Characterization of biotinylated proteins from mature dry pea seeds

Seed coats, cotyledons and embryonic axes were excised from mature dry seeds, and extracts were prepared as described in the Materials and methods section. Figure 3 shows that a major biotinylated polypeptide of molecular mass 65 ± 2 kDa was present in both cotyledons and embryonic axis. This protein is referred to as SBP65, for 65 kDa seed biotinylated protein. On a

per mg of protein basis, the level of SBP65 was nearly the same in both seed compartments. As the total protein content of cotyledons is much higher than that of embryonic axis, the greater proportion of SBP65 is in the cotyledons. In contrast, SBP65 was not detected in the seed coat (Figure 3).

SBP65 was purified from mature dry pea seeds by a two-step chromatographic procedure (Table 1). Proteins specifically eluted from the monomeric-avidin Sepharose column were separated into two groups after chromatography on Mono Q HR5/5 (Figure 4). One group, which did not absorb to this column, comprised the major part of the loaded biotinylated proteins (Figure 4a) and contained SBP65 (Figure 4d). This fraction exhibited no biotin carboxylase activity, using any of the four acceptor substrates, acetyl-CoA, 3-methylcrotonyl-CoA (Figure 4b), propionyl-CoA and pyruvate (not shown). Another group (the salt-eluted fraction) supported MCC and ACC activities (Figure 4b). Again PCC and PC activities were not detectable. SDS/PAGE analysis of this fraction disclosed the presence of four major polypeptides (Figure 4d). The molecular masses of the three fast-migrating polypeptides were 75 kDa, 65 kDa and 50 kDa (Figure 4d). That of the slowly migrating species was
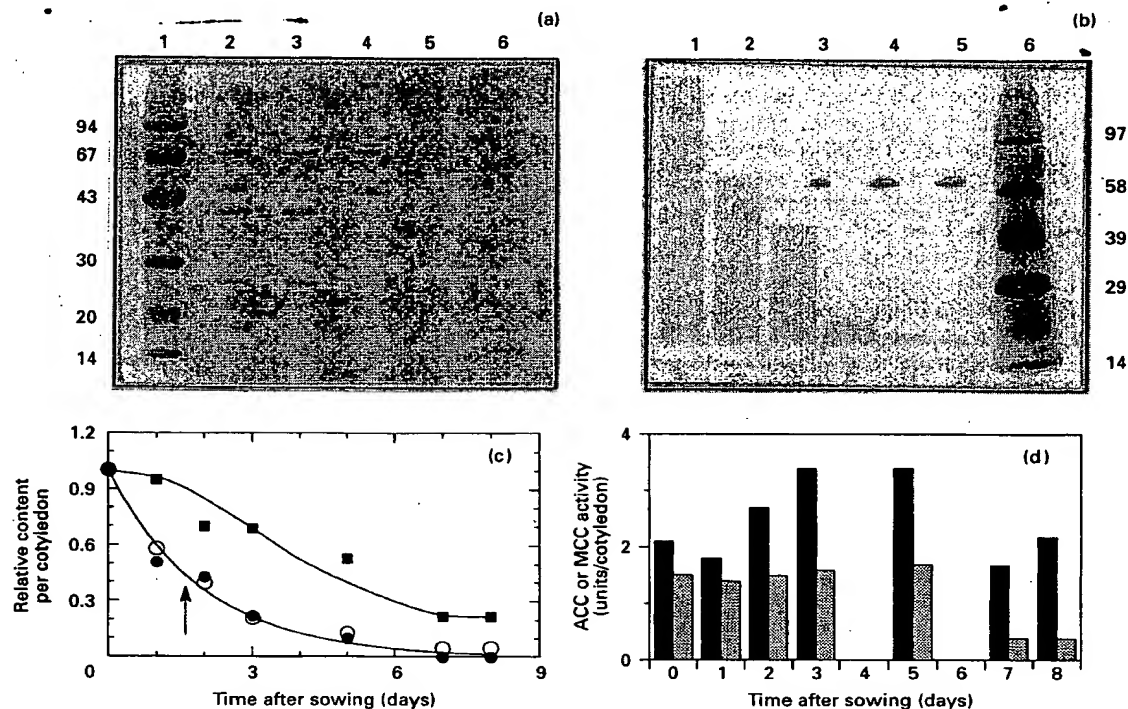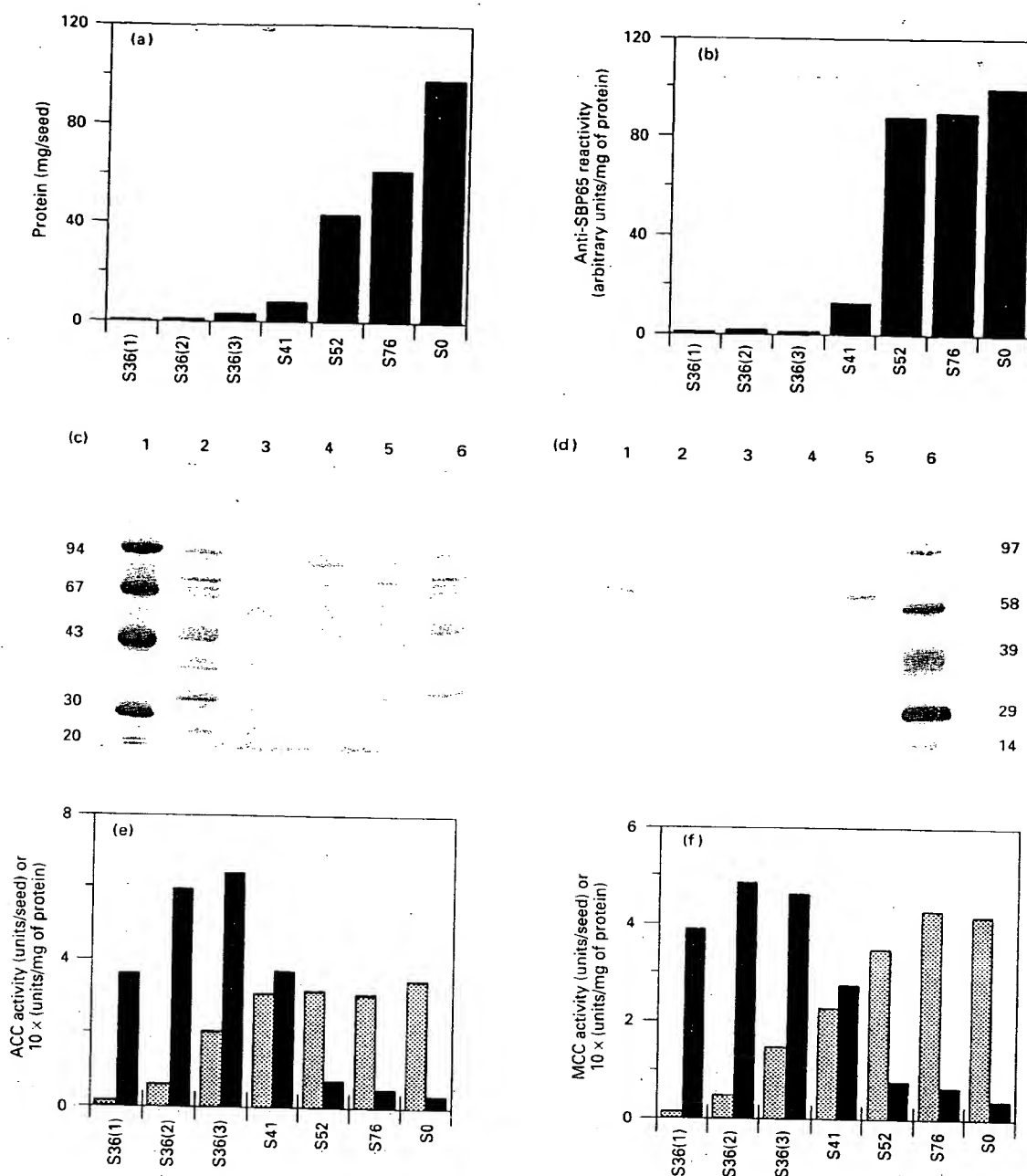
Figure 5   Evaluation of SBP65 in pea seeds and cotyledons during germination

(a) Coomassie Blue R250 staining of a homogeneous SDS/polyacrylamide gel containing 12.5% acrylamide loaded with total protein extracts from mature seeds (S0) (lane 2), seeds collected at day 1 (lane 3) and cotyledons collected at day 3 (lane 4), 5 (lane 5) and 7 (lane 6) after sowing. An amount of 4 μg of protein was loaded in each lane. Molecular-mass markers (lane 1) are indicated in kDa. (b) Nitrocellulose blot of a duplicate of (a) probed with peroxidase-conjugated streptavidin. Lanes 1, 2, 3, 4 and 5 correspond respectively to lanes 6, 5, 4, 3 and 2 in (a). An amount of 16 μg of protein was loaded in each lane. Molecular-mass biotinylated markers (lane 6) are indicated in kDa. (c) Evaluation of total protein (■), protein-bound biotin (○) and anti-SBP65 antibody reactivity (●) in seeds (day zero and 1) and cotyledons (day 2, 3, 5, 7 and 8) at different times after sowing. Radicle emergence is indicated by the arrow. Results are expressed as percentage of corresponding S0 values. (d) Evaluation of ACC (▨) and MCC activities (■).

estimated to be of the order of 200 kDa on SDS/PAGE carried out in gels of higher resolution than that of Figure 4, containing a 7–20% polyacrylamide gradient. Of these four polypeptides, a Western blot analysis with peroxidase-conjugated streptavidin revealed that three (the 65 kDa, 75 kDa and 200 kDa proteins) were biotinylated (not shown). The 65 kDa polypeptide corresponded presumably to contaminating SBP65, which would account for the slight reactivity of the anti-SBP65 antibodies with the salt-eluted fraction (Figure 4a). The 75 kDa biotinylated protein and the 50 kDa non-biotinylated protein could be assigned to the two non-equivalent subunits of MCC. Indeed, this latter enzyme was purified to homogeneity from pea leaves with similar molecular-mass subunits [10]. The 200 kDa polypeptide corresponded presumably to ACC, as this enzyme was purified from pea embryos with similar molecular mass [27]. We note that the purification protocol used by Bettey et al. [27] included a fractionation step on DEAE-Sepharose, and that only the adsorbed proteins were analysed further for purification. Presumably, SBP65 was not detected in the above cited experiments, because the protein adsorbs to neither Mono Q HR5/5 (Figure 4) nor DEAE-TSK (Merck) (not shown). The fact that the 200 kDa and 75 kDa polypeptides were almost not detected in crude extracts (Figure 3) confirmed that SBP65 is the major biotinylated protein in mature dry pea seeds. The purification procedure yielded about 600 μg of SBP65, starting from 100 g of mature dry pea seeds (Table 1).

Biochemical properties of SBP65

An acid hydrolysate [12] of SBP65, but not a native sample of this protein, supported *bioB105* bacterial growth (this strain requires added biotin for growth [28]), suggesting that biotin was covalently bound to the protein. This finding was further supported by the following observations. (i) SBP65 was specifically detected on nitrocellulose filters of blotted SDS/polyacrylamide gels with peroxidase-conjugated streptavidin (Figure 3); (ii) there was no binding or exchange of biotin after incubation of extensively dialysed SBP65 with D-[^{14}C]biotin in the temperature range 25–70 °C, and in the pH range 4–9; and (iii) SBP65 reacted positively and specifically in e.l.i.s.a. conducted with the anti-biotin peroxidase-conjugated antibodies. When SBP65 was allowed to react with CNBr (which cleaves polypeptide chains on the carboxyl side of methionine residues) and the resulting mixture assayed for protein-bound biotin, the reacted sample contained less than 0.5% of the biotin present in the native SBP65 control. An indirect solid-phase biotin assay showed that biotin was released in the supernatant when the treated SBP65 was precipitated by cold acetone. The same behaviour was observed with a sample of biotin–carboxylase mixture, as obtained from the Mono Q column. As the consensus sequence for covalent binding of biotin to biotin carboxylases is Met-Lys-Met, where the vitamin is attached to the lysine residue [2,29], these results suggest a similar mode for binding of biotin

Figure 6   Evaluation of SBP65 during seed formation and maturation

Crude extracts were prepared from various seed samples as described in the Materials and methods section. See the Materials and methods section for nomenclature for samples. (a) Evaluation of total protein. (b) Evaluation of anti-SBP65 antibody reactivity. (c) Coomassie Blue R250 staining of a homogeneous SDS/polyacrylamide gel containing 12.5% acrylamide loaded with 4 μg of total protein extracts from mature dry seeds (S0) (lane 2), and S36(3) (lane 3), S41 (lane 4), S52 (lane 5) and S76 (lane 6) seeds collected from pods. Molecular-mass markers (lane 1) are indicated in kDa. (d) Nitrocellulose blot of a duplicate of (c) probed with peroxidase-conjugated streptavidin. An amount of 16 μg of protein was loaded in each lane. Lanes 1–5 correspond to S76, S52, S41, S36(3) and S0 seed samples respectively. Molecular-mass biotinylated markers (lane 6) are indicated in kDa. (e) Evaluation of ACC activity on a per seed (▨) or on a per mg of protein basis (■). (f) Evaluation of MCC activity on a per seed (▨) or on a per mg of protein basis (■).

to SBP65. The biotin content of SBP65, as determined using the indirect solid-phase biotin assay, yielded a stoichiometry of 0.95 ± 0.15 mol of biotin per 1 mol of 65 kDa polypeptide. The native molecular mass of SBP65 was estimated by gel filtration on Sephacryl S-300 HR to be 450 ± 60 kDa. This indicates that the native protein assumes a polymeric structure, resulting from the association of six to eight identical subunits.

Developmental patterns of SBP65 expression in pea seeds

Because SBP65 appears to correspond to the major biotinylated protein in pea seeds, one might expect similar developmental patterns for the expression of this protein and for that of total protein-bound biotin. In agreement with this expectation, there was a rapid fall for SBP65 from seeds and cotyledons in early

...ges of germination (Figure 5b), which was even more rapid than that of total extractable proteins (Figure 5c). It is known that, in *Pisum sativum*, mobilization of protein reserves from the cotyledons commences just after radicle emergence [30]. SBP65 was absent in the initial period of seed development, and then accumulated considerably in later stages. This accumulation paralleled that of storage proteins (Figures 6a, 6b, and 6c). Clearly, the development patterns of ACC and MCC activities differed from that of SBP65 (Figure 5d and Figures 6e and 6f). In the developing seeds both ACC and MCC appeared to be constitutively expressed, being present at all developmental stages. Also, the Western-blot analyses of Figures 5(b) and 6(d) confirmed that the two biotin-dependent carboxylases were only present in much smaller amounts that SBP65 in the final period of seed maturation.

Organ-specific expression of SBP65

Leaves, roots, stems, pods and flowers were harvested from plants from day 4 to 40 after sowing, and crude extracts were prepared as described in the Materials and methods section. In contrast with the results obtained with the mature seeds, SBP65 was not detectable in these extracts, after either incubation of nitrocellulose filters of blotted SDS/polyacrylamide gels with peroxidase-conjugated streptavidin or by conducting e.l.i.s.a. with the anti-SBP65 antibodies. The expression of SBP65 is therefore confined exclusively to the seeds.

DISCUSSION

A major rationale of this work was to characterize the free and protein-bound biotin contents of pea seeds, and to investigate whether the temporal expression of these contents can be used to define developmental stages during seed formation and germination. Ion-exchange chromatography on Mono Q HR5/5 allows one to separate the biotinylated proteins of mature dry pea seeds into two classes. The first one is not adsorbed by this column, and is associated with a single protein which we have called SBP65. This protein is seed specific, being absent from leaves, roots, stems, pods and flowers. No other biotin-containing protein, exhibiting such a strict organ-specific expression, has yet been reported in plants. A second class of proteins is retained by the Mono Q column, and is mainly associated with MCC and ACC activities. In sharp contrast with SBP65, these two biotinylated enzymes do not exhibit specific organ expression, being detected at all developmental stages in all tissues examined.

We have analysed the relative levels of free and protein-bound biotin in various organs of pea plants. For all proliferating tissues, including the young developing seeds, free biotin is in excess of protein-bound biotin. Presumably, in such tissues free biotin plays a specialized role in growth, and biotinylation of the biotin-dependent carboxylases can only be achieved with an excess of the free vitamin. Studies with mammalian cells suggested that biotin might be involved in regulation of diverse processes such as DNA transcription and replication, cell growth and differentiation, highlighting functions for the vitamin other than as a prosthetic group for the biotinylated carboxylases [2]. In marked contrast, only for the mature dry pea seeds is the level of bound biotin greater than that of free biotin, owing to the accumulation of SBP65.

In considering possible roles for SBP65 in seed development, it is plausible that kinetic competition between different protein acceptors for biotin may divert the normal flux of the vitamin from the housekeeping biotinylated carboxylases to SBP65. In

favour of this hypothesis, the results of CNBr cleavage experiments suggest a similar mode for binding of biotin to the pea seed biotinylated carboxylases and to SBP65. Also, SBP65 might constitute a storage form of the vitamin, reserved for the embryo to start its growth during the germination process. It is worth noting in this context the very fast decrease in the level of SBP65 after seed planting. Other possibilities are that SBP65 supports some as yet unidentified enzyme activity, or that it has a regulatory function at the level of gene expression. As the sequences of several biotinylated proteins are now available [31–34], more definitive evidence about the function of SBP65 can be acquired by cloning and determination of the nucleotide sequence of the gene for this protein. An interesting parallel to our observations with the pea seeds is found in a recent study showing the accumulation of a 62 kDa biotin-containing polypeptide during developing of carrot somatic embryos [11]. If the temporal pattern of expression of major biotinylated polypeptides in various seeds were similar, further characterization of pea SBP65 may provide valuable molecular insights for understanding the role of biotin and protein biotinylation in seed development and germination.

REFERENCES

1   Knowles, J. R. (1989) Annu. Rev. Biochem. **58**, 195–221
2   Dakshinamurti, K. and Chauhan, J. (1989) Vitam. Horm. **45**, 337–384
3   Wurtele, E. S. and Nikolau, B. J. (1990) Arch. Biochem. Biophys. **278**, 179–186
4   Hoppe, H. H. and Zacher, H. (1985) Pestic. Biochem. Physiol. **24**, 298–305
5   Burton, J. D., Gronwald, J. W., Somers, D. A., Connelly, J. A., Gengenbach, B. G. and Wyse, D. L. (1987) Biochem. Biophys. Res. Commun. **148**, 1039–1044
6   Burton, J. D., Gronwald, J. W., Somers, D. A., Gengenbach, B. G. and Wyse, D. L. (1987) Pestic. Biochem. Physiol. **34**, 76–85
7   Somers, D. A., Keith, R. A., Egli, M. A., Marshall, L. C., Gengenbach, B. G., Gronwald, J. W.. and Wyse, D. L. (1993) Plant Physiol. **101**, 1097–1101
8   Stumpf, P. K. (1980) in The Biochemistry of Plants (Stumpf, P. K. and Conn, E. E., eds.), vol. 4, pp. 177–204, Academic Press, New York
9   Harwood, J. L. (1988) Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**, 101–138
10 Alban, C., Baldet, P., Axiotis, S. and Douce, R. (1993) Plant Physiol. **102**, 957–965
11 Wurtele, E. S. and Nikolau, B. J. (1992) Plant Physiol. **99**, 1699–1703
12 Baldet, P., Alban, C., Axiotis, S. and Douce, R. (1993) Arch. Biochem. Biophys. **303**, 67–73
13 Shellhammer, J. and Meinke, D. (1990) Plant Physiol. **93**, 1162–1167
14 Schneider, T., Dinkins, R., Robinson, K., Shellhammer, J. and Meinke, D. W. (1989) Dev. Biol. **131**, 161–167
15 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
16 Green, N. M. (1990) Methods Enzymol. **184**, 51–67
17 Bayer, E. A., Ben-Hur, H. and Wilchek, M. (1990) Methods Enzymol. **184**, 217–223
18 Henrikson, K. P., Allen, S. H. G. and Maloy, W. L. (1979) Anal. Biochem. **94**, 366–370
19 Kohansky, R. A. and Lane, D. L. (1985) J. Biol. Chem. **260**, 5014–5025
20 Kohansky, R. A. and Lane, D. L. (1990) Methods Enzymol. **184**, 194–200
21 Glazer, A. N., DeLange, R. J. and Sigman, D. S. (1975) in Chemical Modifications of Proteins. Selected Methods and Analytical Procedures (Work, T. S. and Work, E., eds.), pp. 68–120, North-Holland Publishing Company, Amsterdam
22 Sakakibara, H., Watanabe, M., Hase, T. and Sugiyama, T. (1991) J. Biol. Chem. **266**, 2028–2035
23 Nikolau, B. J., Wurtele, E. S. and Stumpf, P. K. (1985) Anal. Biochem. **149**, 448–453
24 Meslar, H. W., Camper, S. A. and White, H. B., III (1978) J. Biol. Chem. **253**, 6979–6982
25 Scheiner, J. and De Ritter, E. (1975) J. Agric. Food Chem. **23**, 1157–1162
26 Motel, A., Günther, S., Clauss, M., Kobek, K., Focke, M. and Lichtenthaler, H. K. (1993) Z. Naturforsch. **48c**, 294–300
27 Bettey, M., Ireland, R. J. and Smith, A. M. (1992) J. Plant Physiol. **140**, 513–520

28   Eisenberg, M. A. (1975) J. Bacteriol. **123**, 248–254
29   Dakshinamurti, K. and Chauhan, J. (1990) Methods Enzymol. **184**, 93–102
30   Bewley, J. D. and Black, M. (1983) in Physiology and Biochemistry of Seeds in Relation to Germination, vol. 1, pp. 177–244, Springer-Verlag, Berlin
31   Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C. and Wood, H. G. (1988) J. Biol. Chem. **263**, 6461–6464

32   Kondo, H., Shiratsuchi, K., Yoshimoto, T., Masuda, T., Kitazono, A., Tsuru, D., Anai, M., Sekiguchi, M. and Tanabe, T. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 9730–9733
33   Al-Feel, W., Chirala, S. S. and Wakil, S. J. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 4534–4538
34   Li, S. J., Rock, C. O. and Cronan, J. E., Jr (1992) J. Bacteriol. **174**, 5755–5757

Purification and Properties of the Biotin Repressor

A BIFUNCTIONAL PROTEIN*

Max A. Eisenberg‡, Om Prakash§, and Shu-Chi Hsiung

From the ‡Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, New York 10032 and the §Department of Molecular Virology, Sloan Kettering Institute for Cancer Research, New York, New York 10021

Definitive evidence is presented for the bifunctional nature of the biotin repressor protein which possesses both regulatory and enzymatic activities. The repressor protein can activate biotin in the presence of ATP to form biotinyl-5′-adenylate, the co-repressor which remains tightly bound to the repressor protein. This complex can either bind to the operator site and inhibit transcription or transfer the biotinyl moiety to a lysine residue of the apoenzyme of acetyl-CoA carboxylase. The two activities were coincident throughout a purification procedure which resulted in a 3500-fold increase in activity. Gel electrophoresis of the purified preparation, under native or denaturing conditions, showed three proteins with the activity corresponding to the major protein band of apparent $M_r = 34,000$. On gel exclusion chromatography, the activity was also associated with a protein of M_r varying from 37,000–44,000, indicating the protein is monomeric. The occasional appearance of multiple bands with biological activity in the native gels suggests that the repressor protein can also exist in multimeric forms. On chromatofocusing, the repressor activity and the holoenzyme synthetase activity were coincidental, with the peak of activity at pH 7.2, the isoelectric point. Only a single protein band with $M_r = 34,000$ was observed on SDS gel electrophoresis of all fractions showing activity.

Regulatory mutants of the biotin operon, *bioR*, have been isolated by Eisenberg (1) on the basis of their resistance to the biotin analogue, α-dehydrobiotin, a growth inhibitor. The enzymes of the biotin operon are fully derepressed in these mutants, which also exhibit high biotin excretion. Similar mutants have also been isolated by Pai (2) on the basis of the latter property. Co-transductional analysis by both groups has placed *bioR* at min 89, near the *bfe* locus, on the *Escherichia coli* genetic map. A rifampicin-dominant transducing phage, λ*drif*[d] isolated by Kirschbaum and Konrad (3), has been found by transductional analysis to carry the *bioR* gene and has been used to amplify the *bioR* gene product, the biotin repressor protein. With an *in vitro* protein-synthesizing system, Prakash and Eisenberg (4) have been able to demonstrate the repression of the synthesis of two of the biotin biosynthetic enzymes when the system was fortified with both the partially purified repressor protein and biotin. In a subsequent study of the interaction of the repressor protein-biotin complex with the operator site of the biotin operon, these investigators concluded that the true co-repressor is biotinyl-5′-adenylate. The reason that biotin could function in the *in vitro* system was because the partially purified repressor protein preparation could also activate biotin in the presence of ATP to form biotinyl-5′-adenylate (5). No definitive conclusion could be reached whether the repressor activity and the enzymatic activity are properties of a single protein.

A *bir* mutant, isolated by Campbell *et al.* (6), appears similar to the *bioR* mutants in that the enzymes of the biotin operon are derepressed. This mutant, in contrast, requires high biotin concentrations for growth and shows decreased permeability to biotin. The suggestion was made that a single protein may be responsible for this pleiotropic phenotype. Although initial mapping placed the mutation near *thi*, a subsequent analysis by Pai and Yau (7) indicated that *bir* and *bioR* probably map at the same locus. In an extensive biochemical and genetic analysis, Barker and co-workers (8–10) have determined that *bir* is the structural gene for the enzyme, biotin holoenzyme synthetase. This enzyme activates biotin in the presence of ATP and links it covalently to a lysine moiety of acetyl-CoA carboxylase, the predominant biotin enzyme of *E. coli*. Complementation analysis disclosed the complete or nearly complete overlap of the *bir* and *bir*R genes, suggesting their possible identity. In addition, these investigators have been able to demonstrate that a partially purified preparation of the holoenzyme synthetase protein would protect the operator site from the *Taq*I restriction enzyme, but only in the presence of biotin-5′-adenylate. Thus, the evidence from both laboratories suggests a bifunctional protein with both regulatory and enzymatic properties.

The present study reports the isolation of a homogeneous repressor protein. We have demonstrated that a single protein possesses both regulatory and enzymatic activities: a unique property among repressor proteins thus far isolated.

MATERIALS AND METHODS[1]

RESULTS

A summary of the purification procedure, shown in Table I, indicates about a 3500-fold purification of the repressor protein. This actually represents a 24,000–30,000 increase in

[1] Portions of this paper (including "Materials and Methods" and Figs. 1–6) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 82M-1623, cite authors, and include a check or money order for $4.40 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.
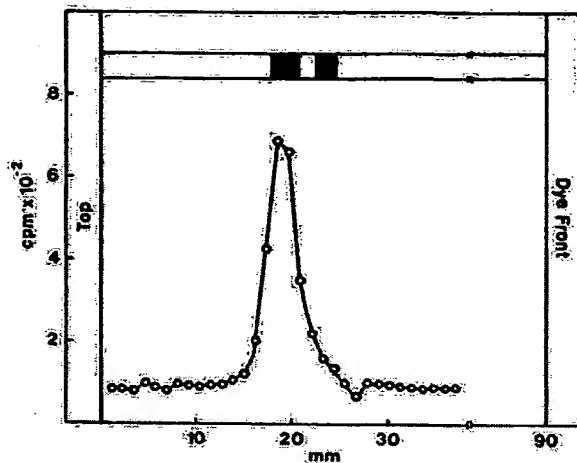
FIG. 9. Polyacrylamide slab gel electrophoresis and the biotin activation profile of fraction 16. Protein samples, 7 and 28 µg, were placed in adjacent wells of a 7% polyacrylamide slab gel. On completion of the run, the gel was cut lengthwise to remove the entire area under the well containing the higher protein concentration. This was sliced into 1.15-mm sections which were assayed for biotin activation activity. The remaining gel was stained with Coomassie blue.
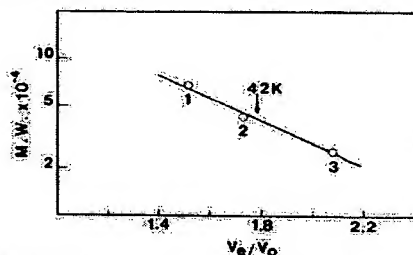

FIG. 10. Molecular weight determination by gel exclusion chromatography. One ml of the 60% ammonium sulfate fraction (biotin-binding activity of 302 pmol/ml) was chromatographed on a Sephacryl S-200 column (1.6 × 92 cm). The elution buffer was 0.01 M sodium phosphate, pH 6.5, containing 0.5 M NaCl, 5% glycerol, 1 × 10 M dithiothreitol. The biotin binding activity and A_{280} were determined on each fraction. The column was previously equilibrated with the following molecular weight (M.W.) protein markers 1, bovine serum albumin, 66,200; 2, ovalbumin, 45,000; 3, chymotrypsinogen, 25,700. The arrow indicates a V_e/V_0 of 1.78 for the repressor protein with $M_r = 42,000$.
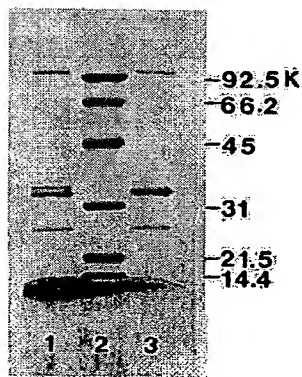
FIG. 11. SDS-polyacrylamide slab gel electrophoresis. Lane 1, DNA-cellulose II chromatography, 6.6 µg of protein; lane 2, SDS-polyacrylamide gel electrophoresis low molecular weight protein markers; lane 3, fraction 16 of DEAE-cellulose II chromatography, pH 6.5, 3.5 µg of protein.

tography reduced the number of protein bands to three, corresponding to $M_r = 98,000$, 26,000, and 34,000 in the order of increasing staining intensity.

While the high molecular weight component could be excluded on the basis of the results from gel exclusion chromatography, it was essential to determine which of the two remaining proteins was associated with the biological activity. Two-dimensional electrophoresis was therefore carried out with the most purified fraction. The native gel was sliced, and each slice was prepared for SDS electrophoresis according to the method of O'Farrell (19) and placed into a separate well of an SDS gel. After staining the gel, only the $M_r = 98,000$ and 34,000 proteins were visible. The migration of the latter protein appeared to be coincident with the distribution curve for the biotin-binding activity obtained on a companion native gel (data not shown).

Chromatofocusing—Although the two-dimensional electrophoresis strongly suggested that the biological activity was associated with the $M_r = 34,000$ protein, it was essential to obtain a more homogeneous protein preparation to determine if both activities resided in the same protein. Chromatofocusing over the pH range of 6.0–8.0 proved to be most successful. In Fig. 12 are shown the results of a chromatofocusing chromatogram on another repressor preparation with somewhat lower specific activity. This preparation had also been purified through the DEAE-cellulose II stage but the first DEAE-cellulose step was omitted. It is apparent that biotin binding, *in vitro* repression, and holoenzyme synthetase activity are all coincidental, with an apparent pI of 7.2. When SDS gel electrophoresis was carried out on the active fractions, the results shown in Fig. 13 were obtained. In order to visualize any possible minor protein bands, the gel was overloaded with the original sample which showed three major protein bands corresponding to $M_r = 98,000$, 34,000, and 31,000. The $M_r = 26,000$ protein observed previously in fraction 16 (Fig. 11) was not seen. A minor band of 60,000 daltons was just barely visible. Since the protein content of the fractions was low, aliquots of 0.3 ml were concentrated under vacuum to a small volume prior to treatment with the SDS buffer. The broadening observed in the protein bands may be due to the
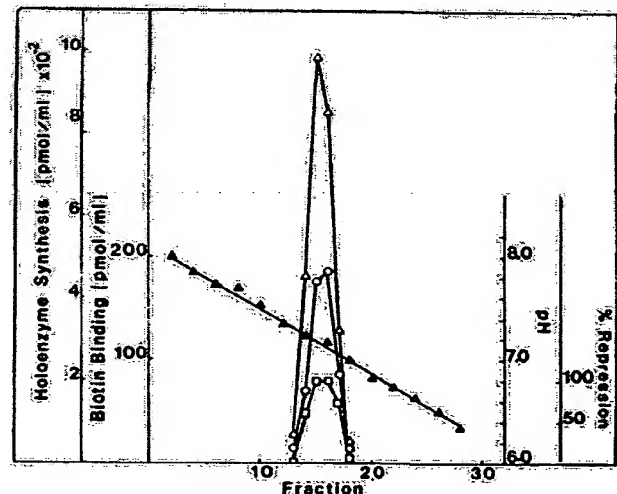

FIG. 12. Chromatofocusing. The procedure was as outlined under "Materials and Methods." One ml of a DEAE-cellulose II preparation containing 290 µg of protein was applied to a 5-ml column and eluted with polybuffer 96-acetic acid which established a pH gradient from 8–6. Biotin-binding activity (○), holoenzyme synthetase activity (△), repressor activity (□), pH gradient (▲).

239, 2858–2864
15. Davis, B. J. (1964) *Ann. N. Y. Acad. Sci.* **121**, 404–427
16. Laemmli, U. K. (1970) *Nature (Lond.)* **227**, 680–685
17. Layne, E. (1975) *Methods Enzymol.* **3**, 448–450
18. Bradford, M. (1976) *Anal. Biochem.* **72**, 248–254
19. O'Farrell, P. H. (1978) *J. Biol. Chem.* **250**, 4007–4021
20. Pirrota, V., Chadwick, P., and Ptashne, M. (1970) *Nature (Lond.)*

227, 41–44
21. Riggs, A. D., and Bourgeois, S. (1968) *J. Mol. Biol.* **34**, 361–364
22. Eisenberg, M. A. (1973) in *Advances in Enzymology* (Meister, A., ed) Vol. 38, pp. 317–372, Interscience Publishing Inc., New York
23. Yanofsky, C. (1981) *Nature (Lond.)* **289**, 751–758
24. Pai, C. H., and Lichstein, H. C. (1965) *Biochim. Biophys. Acta* **100**, 28–35
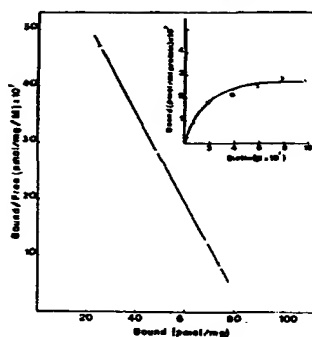
Fig. 1S: Biotin binding as a function of biotin concentration (inset). Scatchard plot (main). Repressor protein concentration was 3.8 mg/ml.
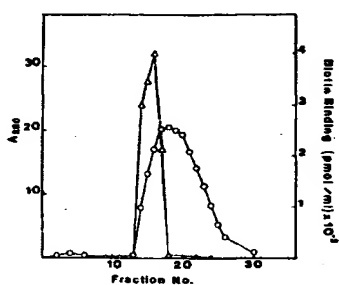
Fig. 2S: DEAE-cellulose I chromatography, pH 7.6 A_{280} (O), biotin binding activity (Δ).
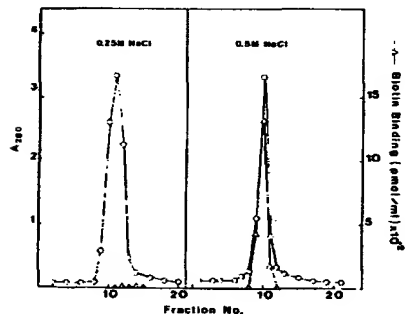
Fig. 3S: Phosphocellulose chromatography, pH 6.5 A_{280} (O), biotin binding activity (Δ).
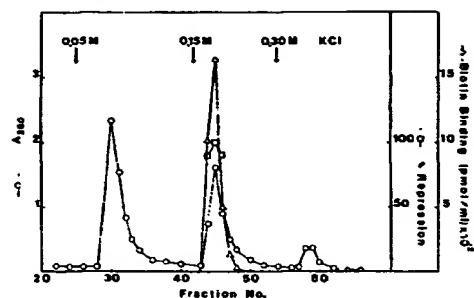
Fig. 4S: DNA-cellulose I chromatography, stepwise elution. A_{280} (O), biotin binding activity (Δ), repressor activity (□).

Fig. 5S: DNA-cellulose II chromatography, gradient elution. A_{280} (O), biotin binding activity (Δ), repressor activity (□).

Fig. 6S: DEAE-cellulose II chromatography, pH 6.5, gradient elution. A_{280} (O), biotin binding activity (Δ), holoenzyme synthetase activity (□), ---NaCl gradient.

AVIDIN

By N. MICHAEL GREEN

National Institute for Medical Research, Mill Hill, London, England

I. INTRODUCTION

A. History

The early history of the protein avidin is part of the story of the discovery, isolation, and synthesis of biotin. Although avidin accounts for only 0.05% of the protein of egg white, its presence was betrayed by an unusual dermatitis in rats fed with dried egg white as the sole source of protein (Boas, 1927). This could be cured by a protective factor, called vitamin H by György, which was present in

many foodstuffs. The history of these investigations in the 1930s and 1940s has been well summarized by György (1954). They led to the identification of vitamin H with biotin (György et al., 1940), the synthesis of biotin, and the partial purification of avidin (Eakin et al., 1941). Although these preparations contained only 30% of active avidin, it proved possible to obtain a crystalline protein from them (Pennington et al., 1942). The activity was decreased by crystallization, and the observation was not followed up. However, in spite of their low activity the crystals appear to be isomorphous with those of pure avidin obtained many years later (Green and Toms, 1970) under rather similar conditions of crystallization.

Since that time interest in the nutritional effects of avidin has waned, but its biotin binding properties have continued to attract the attention of protein chemists and enzymologists. The dissociation constant of the avidin–biotin complex was clearly very low since avidin inhibited the growth of microorganisms that require biotin concentrations of only 10^{-10} M to 10^{-11} M (Wright and Skeggs, 1944; Hertz, 1946). To obtain insight into the nature of this very firm binding, Fraenkel-Conrat et al., (1952a,b) purified avidin more extensively and studied the effects of chemical modification on the activity. Although the activity was still only 70% that of later preparations, the close agreement of its amino acid composition with that subsequently determined suggests that most of the impurity was either avidin–biotin complex or some other form of inactivated avidin.

Little further interest was taken in avidin until the discovery of the coenzyme function of covalently bound biotin (Wakil et al., 1958; Lynen et al., 1959) made it clear that avidin could be a useful tool in the characterization of an important new class of enzymes. A more detailed study of avidin was commenced at this time in the author's laboratory (Green, 1963a,b,c; Melamed and Green, 1963).

Earlier reviews have dealt with avidin in relation to biotin (György, 1954) and biotinyl enzymes (Knappe, 1970; Moss and Lane, 1971). In this review the chemistry and binding properties will be discussed in more detail.

B. Occurrence and Function

Avidin activity has been observed in the eggs and oviducts of many species of birds and in the egg jelly of frogs, at a maximum concentration of about 0.05% of the total protein. The search for avidin in other tissues has not been extensive, but negative results have been reported for avian and amphibian intestinal mucosa and for mammalian oviduct (Hertz, 1946). The recent discovery of an avidinlike protein in culture filtrates of several species of *Streptomyces* (Chaiet et al., 1963) was totally unexpected. It was observed during screening for antibiotics that these culture filtrates owed their activity against gram-negative bacteria to a synergistic action of high and of low molecular weight components. The antibacterial effects could be reversed by high concentrations of biotin in the medium, and it was shown that the high molecular weight component (MSD 235L, streptavidin) was a biotin-binding protein whereas the low molecular weight one (MSD 235 S, stravidin) inhibited biotin synthesis. The physical and chemical characteristics of the streptavidin proved to be remarkably similar to those of egg white avidin. The structure of stravidin has been determined by Baggaley et al. (1969).

The question of the biological function of biotin binding proteins is frequently raised and has not yet received a clear answer. A storage function is unlikely in view of the occurrence of these proteins in an essentially biotin-free state, although it must be admitted that if avidin were to exist with biotin already bound, the complex would not be detected by methods of assay for avidin. A general catalytic function is unlikely in view of the restricted distribution of the protein. Moreover, the catalytically active region of the biotin molecule is the part which interacts most strongly with the avidin so that it is unlikely to be available for catalysis. An antibacterial function is the most plausible one, but the evidence for it is only circumstantial. Egg white contains several antibacterial proteins (lysozyme, a flavin-binding protein, and conalbumin), and one more would not be surprising. The discovery of streptavidin is perhaps the most convincing support for this view in the light of the great versatility of streptomycetes in devising antibiotic systems.

II. Purification and Covalent Chemistry

A. Purification

The earliest isolation by Eakin et al. (1941) made use of selective solubilization of avidin by dilute salt, from alcohol-precipitated egg proteins. An extension of this method by Dhyse (1954) appears to be the source of most partially purified commercial avidin. The best products contained about 50% of active avidin (7 units/mg). [One unit of avidin binds 1 μg of biotin (4.1 nmoles). Methods of assay are discussed in Section VI,A.] Fraenkel-Conrat et al. (1952a) developed an alternative approach, used earlier for the purification of lysozyme (Alderton et al., 1945), in which egg proteins were adsorbed on bentonite. Avidin was eluted with 1 M dipotassium phos-

phate and purified by ammonium sulfate fractionation to a specific activity of 8–10 units/mg. The product was obtained in three forms, a basic glycoprotein and complexes of this with an acidic glycoprotein (water soluble) and with low molecular weight fragments of DNA (water insoluble). Both of the complexes were dissociable by electrophoresis, and although evidence was produced for some specificity of interaction between the components, it seems most likely that the complexes were artifacts of the isolation procedure. They have not been observed during subsequent purifications. The nature of the low molecular weight DNA component of egg-white (approximately 1 mg per egg) does not appear to have been further investigated.

The unusually high isoelectric point of avidin and the availability of cellulose ion exchangers led logically to the development of an improved procedure based on adsorption of the basic proteins on CM-cellulose at high pH, followed by elution of avidin with ammonium carbonate. Further chromatography on CM-cellulose (Melamed and Green, 1963) gave a product which was later crystallized at pH 5 from ammonium sulfate (3 M) or potassium phosphate (3 M) (Green and Toms, 1970; Green and Joynson, 1970). This method is convenient for large-scale preparation. Affinity chromatographic methods using biotinyl cellulose (McCormick, 1965) or biocytinyl Sepharose (Cuatrecasas and Wilchek, 1968) have been described, but they have not yet been widely used.

B. Homogeneity

Avidin purified on carboxymethyl cellulose and then crystallized appears to contain two or three components when chromatographed on Amberlite CG-50 (Melamed and Green, 1963; Green and Toms, 1970). There were indications of differences of a few amino acid residues between the components but the variable amino acid has differed with different avidin preparations, and the analyses have not been repeated often enough to establish these small differences with certainty. There was some evidence that eggs derived from a single breed of hen gave a more homogeneous product (Melamed and Green, 1963). However, the amino acid sequence of avidin (Section II,D) was determined with protein purified from a commercial source of crude avidin (De Lange, 1970) probably derived from different breeds from those used by Melamed and Green, and the products from the two laboratories had almost identical amino acid compositions (see Table I). The sequence showed only one ambiguity; residue 34 was either threonine or isoleucine (De Lange and Huang,

1971). Since the variant peptides were recovered in equal amounts it is possible that each avidin molecule contains two of each type of chain, though the presence of heterogeneity suggests that the two chains are more likely to originate from different molecules. Small variations in carbohydrate composition have been observed (Huang and De Lange, 1971, and Table I), but further work is required to determine their extent.

C. General Properties and Stability

Avidin is a basic glycoprotein with its isoelectric point at pH 10 (Woolley and Longsworth, 1942). It is very soluble in water and salt solutions, and it is stable over a wide range of pH and temperature. It crystallizes from strong salt solutions between pH 5 and pH 7, but it has not yet been crystallized in the isoelectric region. In contrast, streptavidin has no carbohydrate, it has an acid isoelectric point, it is much less soluble in water and can be crystallized from water or 50% isopropanol (Chaiet et al., 1963).

The early determinations of molecular weight (Woolley and Longsworth, 1942) and of composition (Fraenkel-Conrat et al., 1952b) are surprisingly close to the results obtained with pure avidin, considering that the preparations used possessed only 50–70% of the maximum biotin binding capacity. It is likely that much of the impurity in these earlier preparations was avidin–biotin complex (Pai and Lichstein, 1964) or otherwise inactivated avidin. This might also explain why the crystals obtained by Pennington et al. (1942), which are similar in appearance to crystals of avidin, possessed a lower specific activity than the solution from which they were crystallized. The avidin–biotin complex is isomorphous with avidin (Green and Joynson, 1970) and could well crystallize more readily.

The biotin-binding capacities and amino acid compositions of avidin and streptavidin are compared in Table I. With the exception of the differences just mentioned there are marked similarities in composition. There are zero to two residues each of histidine, proline, cystine, and methionine, and the contents of tryptophan and threonine are very high, that of threonine uniquely so. Other properties of streptavidin, most of them similar to those of avidin, are summarized in Section VI,D.

The remarkable stability of avidin, and even more of the avidin–biotin complex, to heat and to proteolytic enzymes was noted very early. In a study of conditions required to release biotin from the complex by autoclaving it was observed that release was more rapid

Ross, 1973) has shown that avidin is inactivated in an endothermic transition at 85°C, whereas a similar transition with the avidin–biotin complex does not occur until a temperature of 132°C is reached (Section V,B). Loss of tertiary structure at 85°C has also been observed by measurement of the optical rotatory dispersion (ORD) spectrum (Pritchard et al., 1966).

The avidin–biotin complex is resistant to proteolysis by the enzymes of the digestive tract and free avidin is not inactivated by trypsin or by Pronase. However, an early observation of György and Rose (1941) showed that biotin was released from the complex when it was administered intravenously or parenterally. This has been recently confirmed (Wei and Wright, 1964; Lee et al., 1973), but nothing is known of the enzymes responsible. The breakdown is rather slow, and in short-term experiments Miller and Tausig (1964) showed that both avidin and streptavidin administered parenterally could protect mice against Salmonella infection in the presence of stravidin, an inhibitor of biotin synthesis.

The biotin-binding activity is not lost between pH values of 2 and 13 nor at concentrations of guanidinium chloride (GuHCl) below 3 M. Loss of activity outside these limits is accompanied by unfolding of the protein and dissociation into subunits, which is largely reversible (Green, 1963c, and Section V,A).

The instability of avidin under oxidizing conditions, particularly in strong light (György et al., 1942), is probably the result of an unusually reactive tryptophan residue in the binding site (Section II,E).

D. Primary Structure

The complete sequence, established by De Lange and Huang (1971), is shown in Table II, aligned with that of hen egg white lysozyme. It had been suggested earlier (Green, 1968) that the two proteins might have arisen from a common ancestor on the rather tenuous evidence of similar amino acid composition and a weak binding of biotin by lysozyme. The alignment from the C-terminal end shows 15 identities (12%), which is little more than would be expected on a random basis. This observation that 88% of the residues are different is hardly convincing evidence even for a remote relationship. However, it is not very different from the 86% of differences observed between trypsinogen and the α-lytic protease of Sorangium (Dayhoff, 1972), where evidence for homology is made more convincing by the resemblance between the active site sequences. A further point against a chance resemblance is the ef-

TABLE I

Amino Acid Composition of Avidin and Streptavidin[a]

	Avidin residues/subunit	Streptavidin residues/subunit
Lysine	9	4
Histidine	1	2
Arginine	8	4
Aspartic acid	15 (14)	12
Threonine	20.5 (19)	19
Serine	9	10
Glutamic acid	10	9
Proline	2	2
Glycine	11	17
Alanine	5	17
Half-cystine	2	0
Valine	7	7
Methionine	2	0
Isoleucine	7.5 (8)	3
Leucine	7	8
Tyrosine	1	6
Phenylalanine	7	2
Tryptophan	4	8
Total residues	128	ca. 130
Amide	16	–
Mannose	4 (5)	0
Glucosamine	3	0
Subunit weight	15,600	14,000
Molecular weight	67,000[b]	60,000[c]
Biotin-bound/subunit 1%	0.97	0.95
E_{288} (1%)	15.4[d]	34[e]
		47[e]
ϵ_{282} (subunit)	24,000	56,000

[a] The amino acid composition of avidin and the molecular weight of the subunit is based on the sequence (De Lange and Huang, 1971). The composition is identical with that given by Green and Toms (1970), except for figures included from the latter reference in parentheses. The results for streptavidin are from Chaiet and Wolf (1964) and Green (1968).
[b] See Table III.
[c] Chaiet et al. (1963).
[d] Green and Toms (1970).
[e] Green (1970).

in the absence of salts and was 88% complete after 10 minutes at 100°; however, autoclaving at 120° for 15 minutes was required for complete release in the presence of salts (Pai and Lichstein, 1964; Wei and Wright, 1964). A recent calorimetric study (Donovan and

How many of the four tryptophans interact directly with bound biotin? (2) Is it possible to modify them selectively? (3) Is any one of them of particularly vital importance? Taking the last questions first, oxidation with N-bromosuccinimide showed that an average of one tryptophan could be destroyed with loss of only 25% of activity, but that all activity disappeared when a second residue was oxidized. Huang (1971) isolated tryptophan peptides after partial inactivation. His preliminary results showed that when half the activity was lost (about 1.3 tryptophans oxidized) about half of tryptophans 10 and 70 were destroyed, Trp 97 was intact, and the evidence on Trp 110 was inconclusive. From this it appears that either or both Trp 10 and Trp 70 are essential to the binding. More selective oxidation could be effected by sodium periodate (1 mM) at neutral pH (Green, 1963b, 1974). It produced a slow fall in E_{280} ($t_{1/2} = 35$ minutes) which stopped after oxidation of about one residue of tryptophan. Neutral periodate has been shown to oxidize indoles and tryptophan derivatives to formyl kynurenine derivatives (Dolby and Booth, 1966), but the rate is less than one-tenth as fast as that for the tryptophan in avidin. The spectrum of the oxidized avidin was consistent with oxidation of slightly more than one residue to formyl kynurenine. The product possessed 70% of the initial binding capacity and its affinity for biotin was decreased from 10^{-15} M to about 10^{-9} M. Its main feature of interest was the modified difference spectrum given with biotin, which showed peaks characteristic of N-formylkynurenine (Section IV,C). A product with similar spectroscopic properties has also been obtained by autoxidation of avidin during concentration of solutions by pervaporation (Green, 1963b), but the conditions are not readily reproducible.

To return to the questions posed above, the tryptophan oxidized by periodate does not appear to be of vital importance for biotin-binding though the affinity is significantly reduced. In contrast, avidin in which Trp 10 and Trp 70 were oxidized was inactive, but it is not clear whether this was because the residues were different from that oxidized by periodate or because two residues were attacked rather than one. It may also be significant that oxindole, the major product of oxidation by N-bromosuccinimide, does not contribute to the difference spectrum of the partially inactivated product whereas formylkynurenine does contribute to that of the periodate-oxidized avidin (Green, 1963b, and Section IV,C). Although no compelling arguments can be made from this evidence, it seems likely that several tryptophans contribute to the biotin-binding site. The only other functional group which has been implicated in the

binding site of avidin is an amino group. It was observed that the initial rate of reaction of 1-fluoro-2,4-dinitrobenzene (FDNB) with avidin was faster than that with the avidin–biotin complex and that, if the reaction was stopped after one DNP had been introduced per subunit, almost no biotin binding activity remained (Green, 1975). This suggested that there was a unique amino group in the neighborhood of the binding site with which FDNB reacted preferentially. In contrast trinitrobenzene sulfonic acid and dansyl chloride (Green, 1963c) reacted without blocking this residue. The location of the lysine which is blocked by FDNB has not been determined. It may be one of the three lysine residues (9, 71, 111) which are adjacent to tryptophans in the sequence.

The single tyrosine residue (33) of avidin ionizes only at pH values above 12 and is probably buried (Green, 1963c). In the avidin–biotin complex it remained nonionized even in 0.5 M KOH. The single histidine, His (50), was alkylated by iodoacetamide only in the presence of 6 M guanidinium chloride (Huang, 1971). The single disulfide residue (4–62) could be reduced only after dissociation of the tetramer into subunits (Green, 1963c) and is also likely to be a buried residue. Preliminary attempts to regenerate activity from the reduced subunits (0.5 mg/ml) were unsuccessful (Green, 1963c).

It is unlikely that the carbohydrate plays any part in biotin binding. Most of the mannose residues would have been oxidized by the periodate treatment and the product still binds biotin strongly. Furthermore, streptavidin contains no carbohydrate.

III. PHYSICAL PROPERTIES

A. Hydrodynamic Properties

The hydrodynamic properties of avidin have not been systematically studied. Sedimentation velocity, sedimentation equilibrium, and diffusion coefficient measurements have been made, mostly at single concentrations, but there have been no reports of viscosity measurements. The available results are summarized in Table III together with the molecular weights determined by osmotic pressure, and the rotational relaxation times of dansyl avidin calculated from measurements of polarization of fluorescence. The results obtained by different methods or in different laboratories under different conditions agree well and suggest that the shape and conformation of avidin is not sensitive to changes in temperature, pH, and ionic strength. For example, the sedimentation and diffusion coefficients of avidin at pH 5.5 are the same as those of the avidin–biotin

from which a rotational relaxation time, ρ_h of 56 nsec could be calculated. This was approximately the same as ρ_o, the theoretical relaxation time for an anhydrous spherical protein of the same molecular weight. The ratio ρ_h/ρ_o is usually nearer 2 than 1 for a rigid molecule. This result implies some freedom of rotation within the conjugate. Later measurements performed in sucrose at constant temperature (Fig. 1) (Green, 1964a) gave a relation between $1/p$ and T/η which departed from linearity in viscous solutions (about 40% sucrose), resembling the results of Wahl and Weber (1967) with labeled immunoglobulin. The value of ρ_h (100 nsec) calculated from the linear section of the curves was nearer the expected value for a rigid molecule, while the curvature at high viscosity implies a thermally activated rotation of the dansyl group relative to the protein, which was considerably greater in the avidin–biotin complex than in avidin. This would account for the lower apparent value of ρ_h obtained when temperature was varied. This greater freedom of rota-

FIG. 1. Effect of biotin on the polarization of fluorescence of dansyl avidin. The polarization of fluorescence was measured at 25°C in the presence of increasing concentrations of sucrose, using avidin which carried one dansyl group per subunit, as described by Wahl and Weber (1967). ○, Dansyl avidin; ■, dansyl avidin + biotin.

TABLE III

Physical Properties of Avidin and Avidin–Biotin Complex

Parameter	Avidin		Avidin–biotin	
	Neutral buffers	Denaturing solvents	Neutral buffers	Denaturing solvents
$s_{20,w}$ (Svedberg units)	4.7[a]	0.89 (6 M GuHCl)[c]		4.58 (0.1 M HCl)[d]
$D_{20,w}$ (cm²/sec × 10⁷)	4.55[d]	2.15 (0.1 M HCl)[e]		6.08 (0.1 M HCl)[d]
	5.98[d]	4.5 (6 M GuHCl)[c]		
f/f_0	1.33	3.7 (6 M GuHCl)[c]		
Molecular weight				
Sedimentation and diffusion	66,000[d]			66,000 (0.1 M HCl)[d]
Sedimentation equilibrium	68,000[d]	18,700 (3 M GuHCl)[c]		
Osmotic pressure	69,000[f] 66,000[b]	68,000 (8 M Urea)[b] 17,000 (3 M GuHCl)[c]		
Fluorescence polarization				
ρ_h nsec (T varied)	56[c]	ca 25 (3 M GuHCl)[c]	50[c]	ca 50 (6 M GuHCl)[c]
ρ_h nsec (isothermal)	100[d]		105[d]	

[a] Woolley and Longsworth (1942).
[b] Fraenkel-Conrat et al. (1952a).
[c] Green (1963c).
[d] Green (1964b).
[e] Green and Melamed (1966).
[f] De Lange (1970).

complex in 0.1 M HCl. The observation that crystals of avidin are isomorphous with those of the avidin–biotin complex (Green and Joynson, 1970) provides further evidence that there is little change in structure when biotin is bound. The results of molecular weight measurement are within 5–10% of that expected for a molecule containing four of the chemically defined subunits (62,400, Table I) and four biotin-binding sites (64,800). This is within the limits of error of the physical measurements. The molecule dissociates into four subunits in GuHCl (6 M) or HCl (0.1 M) without reduction of the disulfide bonds. The subunits do not reassociate until the GuHCl concentration is decreased below 3 M (Section V,A).

Measurements of polarization of fluorescence of dansyl avidin as a function of temperature gave a linear relation between $1/p$ and T/η

FIG. 2. Difference spectra of avidin complexes versus avidin. The avidin subunit concentration was 15.3 μM in 50 mM sodium phosphate pH 6.8 (a–e) or 50 mM sodium borate pH 9.2 (f). The ordinate scale of the first derivative spectrum is arbitrary. The measurements were made with a Cary 118 spectrophotometer. (a) Biotin; (b) first derivative of avidin spectrum; (c) biotin methyl ester; (d) desthiobiotin; (e) desthiobiotinol; (f) diamine from biotin.

tion induced by biotin could be correlated with a change in the dansyl fluorescence spectrum, which shifted from 508 nm to 514 nm and lost 40% of its intensity when biotin was bound. This suggests a more aqueous environment for the dansyl groups and implies that in avidin they are located near hydrophobic regions of the binding site and are displaced into solvent by biotin. Unlike the 2,4-dinitrophenyl groups discussed above (Section II,D), the dansyl groups do not block the binding site.

B. Spectroscopic Properties

1. Ultraviolet Absorption

The spectroscopic properties of avidin proved much more sensitive than the molecular kinetic properties to combination with biotin (Table IV), suggesting direct involvement of aromatic amino acid residues in the biotin-binding site (Green, 1962, 1963b). Since each subunit of avidin contains four tryptophan residues and only one tyrosine, the ultraviolet spectrum is dominated by the contribution of tryptophan. The sum of the contributions of tyrosine and tryptophan

accounts for 96% of the absorbance at 282 nm, allowing for a red shift of 2 nm relative to the free amino acids. When biotin binds to avidin there is a further red shift and a narrowing of the 280 nm band leading to the difference spectrum shown in Fig. 2, where it is compared with the difference spectra given by a number of analogs. The first derivative of the avidin spectrum is included to provide a stan-

TABLE IV

Spectroscopic Properties of Avidin

| | Avidin[a] | Avidin–biotin[a] | Avidin In GuHCl | | Reference[b] |
			6.9 M	3.4 M	
Ultraviolet					
ϵ_{282} (subunit)	24,000	24,000	23,000	–	1
ϵ_{233} (subunit)	93,000	118,000	74,000	–	1
$\Delta\epsilon_{294}$	0	1,800	−2,050	−2,400	1
$\Delta\epsilon_{233}$	0	25,000	−14,000	−19,000	1
Circular dichroism					
$(\epsilon_l - \epsilon_r)_{292}$	3.50	5.15	0.3	–	2
$(\epsilon_l - \epsilon_r)_{228}$	49.8	46.0	−2	–	2
Fluorescence					
Emission λ_{max}	338 nm	328 nm	350 nm	–	3
Intensity (tryptophan) = 1.0 at 353 nm)	0.63	0.50	0.90	–	3

[a] Sodium phosphate, 0.05 M, pH 7.
[b] Unpublished measurements on crystalline avidin using a Cary 118 spectrophotometer, based on a subunit weight of 15,600; 2, Unpublished measurements on crystalline avidin using a Cary 61 circular dichroism spectropolarimeter. The values of $(\epsilon_l - \epsilon_r)$ were calculated per mole of tryptophan; 3, Green (1964a).

2. Fluorescence

The fluorescence emission spectrum of avidin (Green, 1964a) is typical of many proteins of high tryptophan content (Teale, 1960). The changes that follow combination with biotin or dissociation into subunits (Table IV) are consistent with the changes in exposure to the solvent deduced from the shifts in absorption spectrum.

3. Circular Dichroism and Secondary Structure

The CD and ORD spectra of avidin and streptavidin, like the absorption and fluorescence spectra, are dominated by the contribution of tryptophan (Green and Melamed, 1966). We shall consider only the CD spectra (Fig. 3 and Table IV) since they contain all the essential information. They are unusual in that they show two positive bands. The longer wavelength band extends from 250 nm to 300 nm and shows fine structure that can be matched closely by the CD spectrum of tryptophan measured at 77°K in methanol/glycerol mixtures (Strickland et al., 1969). In avidin the ^1L_a bands predominate. In the complex with biotin the intensity of the ^1L_a band remains

FIG. 3. Circular dichroism of avidin and avidin–biotin complex. Spectra were obtained with a Cary 61 CD spectropolarimeter [26 μM avidin subunits in 0.1 cm (left-hand scale) or 1-cm cuvettes (right-hand scale)]. (a) Avidin; (b) avidin–biotin complex; (c) base line. $(\epsilon_1 - \epsilon_r)$ is expressed per mole of tryptophan. θ (°/cm²/decimole peptide bond) = 103 $(\epsilon_1 - \epsilon_r)$.

dard showing the effects of a uniform red shift. Comparison of the biotin spectrum with the latter shows a much larger shoulder at 298 nm and a deeper trough at 260 nm than would be expected for a uniform shift. This result is consistent with a larger red shift of the ^1L_a bands of tryptophan (265 nm, 280 nm, and 297 nm) which lie on either side of the ^1L_b bands (at 284 nm and 291 nm) and are often more sensitive to solvent perturbation (Strickland et al., 1969; Andrews and Forster, 1972). In this case, it appears possible to assign the extra perturbation of the ^1L_a bands in part to the negatively charged carboxyl group of biotin and in part to the sulfur atom. Removal of either of these features gives a much closer approach to the derivative spectrum, as can be seen from the spectra given by biotin methyl ester and by desthiobiotin. Removal of both negative charge and sulfur (desthiobiotinol) significantly sharpens the ^1L_b contribution. Whatever the theoretical origin of these features, they provide a useful characteristic of each biotin analog.

The difference peak near 233 nm is an order of magnitude larger than those in the 280 nm region (Table IV). It is much more uniform in shape, consistent with the simpler origins of the 219 nm band of tryptophan, shifted to 226 nm in the protein (Green, 1962, 1963b). The large difference extinction coefficient provides the basis of a spectrophotometric assay for avidin or biotin and of a method for determining dissociation constants of biotin analogs. It decreases as the resemblance of the analog to biotin becomes more remote (Table V). Evidence discussed below (Section IV) suggests that this is related to the number of different tryptophan residues that interact with the analog.

These effects can be understood qualitatively in terms of an exclusion of water from the immediate neighborhood of the tryptophan residues by the bound biotin, together with an additional perturbation from the negative charge affecting mainly ^1L_a bands. This effect of negative charge has also been seen with simple tryptophan derivatives (Ananthanarayanan and Bigelow, 1969).

The far ultraviolet absorption spectrum (Green and Melamed, 1966) is unaffected by combination with biotin, and the mean extinction of coefficient of the peptide chromophore is low ($\epsilon_{196} = 4800$), suggesting a high proportion of α-helix, when interpreted by the criteria of Rosenheck and Doty (1961). This conflicts with the conclusions from CD measurements that there is little or no helix present. A similar unexplained discrepancy exists between far UV and CD spectra of immunoglobulins (Gould et al., 1964).

about the same while that of the ^1L_b bands at 286 nm and 293 nm (2 nm red shift relative to free tryptophan) increase. The high intensity of the bands, four or five times that of the tryptophan derivatives at low temperature or of the aromatic bands of other proteins, implies a rigid asymmetric environment for the chromophores.

The shorter wavelength band (228 nm) showed no fine structure and only a small change of intensity when biotin was bound. Its position is almost exactly that predicted for the red-shifted 219 nm band of tryptophan (Green, 1962). It is unlikely that the peptide chromophore is contributing significantly to the CD in this region since at shorter wavelengths, where its effects should be most readily detectable, the dichroism is very low. In view of this it is difficult to draw any positive conclusions about tertiary structure, but it is probably safe to say that there is little or no α-helix in avidin. Chou and Fasman (1974) have recently proposed an empirical method for calculating the fractions of α-helix and β structure in a protein from the amino acid sequence. Application of their criteria to avidin leads to an estimate of zero α-helix and 50% of β-structure. This would be consistent with the very low CD of avidin below 215 nm, taking into account the averaged spectra calculated by Chen et al. (1972) for a number of proteins. One remaining unknown factor is the contribution of the absorption bands of tryptophan in the region below 215 nm.

In GuHCl (6 M) there was almost no significant dichroism at wavelengths above 215 nm. The ORD spectrum showed that b_o was zero. In 0.1 M HCl the aromatic bands were still significant, showing that unfolding in this solvent was incomplete (Green and Melamed, 1966). This is also implied by the lower mobility of avidin subunits on gel filtration in 0.1 M HCl as compared with 6 M GuHCl (Green and Toms, 1972).

IV. BINDING PROPERTIES

A. Biotin

The early work on avidin showed that the combination with biotin was very firm. Attempts to measure the dissociation constant by equilibrium dialysis with radioactive biotin (Launer and Fraenkel-Conrat, 1951) gave an upper limit of 10^{-10} M; the presence of radioactive impurities in the biotin made more exact interpretation difficult. In later work (Green, 1963a) the dissociation constant was calculated from the ratio of the rate constants for the forward and reverse reac-

tions. The rate of dissociation was determined from the rate of exchange of bound ^{14}C-labeled biotin with an excess of unlabeled biotin, by separating free from bound with CM-cellulose. After a relatively rapid (18-hour) exchange of 5% of the biotin, the remainder of the reaction followed first-order kinetics for as long as it was measured (800 hours, Green and Toms, 1973) with a rate constant of 9×10^{-8} sec^{-1}. The rate was insensitive to ionic strength (N. M. Green, unpublished), but increased when the pH was below 3 (Table V). The rate constant measured by this method is that for the first dissociation step ($AB_4 \rightarrow AB_3 + B$). However, subsequent work (Section IV,E) has shown no interactions between the four binding sites, and it is likely that the calculated dissociation constant applies to all four sites.

The dissociation rate constant of biotin is so low that reequilibration of biotin after the initial binding can be ignored in any but very long-term experiments. This enables the forward rate to be measured by stopping the reaction with biotin-^{14}C by using an excess of unlabeled biotin. The rate constant obtained in this way, 7×10^7 M^{-1} sec^{-1}, was significantly less than that for a diffusion controlled reaction but similar to that observed for a number of enzyme–substrate and antibody–hapten combinations (Hammes, 1969). It should be possible to obtain more accurate values for the forward rate constant by stopped-flow spectrophotometry, but no measurements have been described.

It has been suggested that the stability of the avidin–biotin complex is decreased at very low ionic strength. The evidence is of two kinds. (1) The release of biotin on autoclaving the complex was more effective in the absence of salt (Wei and Wright, 1964; Pai and Lichstein, 1964); and (2) the combination of avidin with biotin appeared to be less complete at low ionic strength, when measured by dialysis or by separation of radioactive avidin–biotin complex by gel filtration (Wei and Wright, 1964). Many attempts have been made in the author's laboratory to show an increased dissociation constant in the absence of salt for complexes with biotin and several analogs, but no effects have been observed, either by following rates of exchange of biotin-^{14}C by measuring rates of dialysis or displacement of thio-biotin ($K = 10^{-13}$ M), or by spectrophotometric titration of more weakly bound analogs. Although the stability of the avidin–biotin complex to heat is less in the absence of salt, it is unlikely that the dissociation constant is changed. A possible source of misleading effects in gel filtration at low ionic strength is considered below (Section VI,B).

TABLE V (Continued)

#		R_2	R_4	R_5	pH	$\dfrac{\Delta E_{max}}{E_{282}}$ [b]		k_{-1} [c] (sec⁻¹)	k_{-1}/k_1 [c] (M)	K [c] (M)	$-\Delta G'$ (kcal/mole)	$-\Delta H'$ (kcal/mole) (6)	Refer-ences [g]
	Imidazolidone derivatives [a]												
8	d-Desthiobiotin	O	(CH₂)₄COOH	CH₃	4.0	0.46		7.6×10^{-5}	10×10^{-12}	—	16.5	—	4
					7.0	0.47		3.6×10^{-5}	5×10^{-12}	—	16.9	—	4
	dl-Desthiobiotin [d]	O	(CH₂)₄COOH	CH₃	4.0	0.46	d	6.3×10^{-5}	9×10^{-12}	—	16.5	—	4
					4.0	—	l	4.3×10^{-5}	6×10^{-11}	—	14.0	—	4
9	dl-Desthiobiotin methyl ester	O	(CH₂)₄COOCH₃	CH₃	7.0	0.39	d	2.9×10^{-4}	4×10^{-12}	—	15.6	—	4
					7.0	—	l	1×10^{-2}	1.5×10^{-11}	—	14.8	—	4
10	dl-Desthiobiotinol [d]	O	(CH₂)₄CH₂OH	CH₃	7.0	0.49	d	3.8×10^{-3}	5×10^{-11}	—	14.1	—	4
					7.0	—	l	7.7×10^{-2}	1×10^{-9}	—	12.3	—	4
11	D-4-n-Hexylimidazolidone	O	(CH₂)₅CH₃	H	7.0	0.38		0.24	3.5×10^{-9}	9×10^{-9}	11.1	—	4
	L-4-n-Hexylimidazolidone	O	(CH₂)₅CH₃	H	7.0	0.36		3.4	5×10^{-8}	2.3×10^{-7}	9.1	—	4
12	dl-4-n-Butylimidazolidone	O	(CH₂)₃CH₃	H	4.0	0.25		—	—	4×10^{-6}	7.4	—	4
					7.0	0.43		—	—	1.3×10^{-6}	8.1	—	4
					10.0	0.48		—	—	1.7×10^{-6}	7.9	—	4
13	dl-4-n-Propylimidazolidone	O	(CH₂)₂CH₃	H	7.0	0.24		—	—	5×10^{-6}	7.3	—	4
14	dl-4-Ethylimidazolidone	O	CH₂CH₃	H	7.0	0.22		—	—	8×10^{-6}	7.0	—	4
15	dl-4-Methylimidazolidone	O	CH₃	H	7.0	0.23		—	—	3.4×10^{-5}	6.1	—	4
16	Imidazolidone	O	H	H	6.8	0.26		—	—	5×10^{-4}	4.5	—	2
17	dl-4,5-Dimethylimidazolidone	O	CH₃	CH₃	7.0	0.17		—	—	2.3×10^{-5}	6.4	—	4
	meso-4,5-Dimethylimidazolidone	O	CH₃	CH₃	7.0	0.19		—	—	3×10^{-6}	7.5	—	4
18	dl-Norleucine hydantoin	O	(CH₂)₃CH₃	O	7.0	0.2		—	—	4×10^{-5}	6.0	—	4
19	D-4-n-Hexyl-2-thionoimidazolidine	S	(CH₂)₅CH₃	H	7.0	0.23		—	—	1.2×10^{-5}	8.1	—	4
20	D-4-n-Hexyl-2-iminoimidazolidine	NH⁺	(CH₂)₅CH₃	H	11.8	0.26		—	—	4×10^{-6}	7.4	—	4
		NH²	(CH₂)₅CH₃	H	Calculated for free base [e]			—	—	1×10^{-6}	8.2	—	4
	Oxazolidone derivatives [a]												
21	D-4-n-Hexyloxazolidone	O	(CH₂)₅CH₃	H	7.0	0.12		—	—	2.2×10^{-4}	5.0	—	4
22	D-5-n-Hexyloxazolidone	O	H	(CH₂)₅CH₃	7.0	0.26		—	—	4.4×10^{-6}	7.4	—	4
	Miscellaneous compounds												
23	Diamine from biotin [5-(3,4-diaminothiophan-2-yl)+ pentanoic acid]	—	—	—	4.6	0.36		—	—	5×10^{-3}	2.7	—	2
					6.8	0.73		—	—	3×10^{-7}	9.0	—	2
					10.0	—		—	—	1.2×10^{-7}	9.5	4.7	6
24	Lipoic acid	—	—	—	6.8	0.39		—	—	7×10^{-7}	8.5	—	2
25	Tetrahydrofuran	⋜	—	—	7.0	0.28		—	—	0.9×10^{-2}	2.8	—	4
26	Urea	—	—	—	6.8	0.33		—	—	3.6×10^{-2}	2.2	—	2
27	Ethylene glycol	—	—	—	6.8	0.2		—	—	0.5	1	—	2
28	Hexyl alcohol	—	—	—	7.0	0.09		—	—	4.4×10^{-3}	3.2	—	4
29	Hexanoic acid	—	—	—	6.8	0.16		—	—	3×10^{-3}	4.9	—	2
						E_{500}/E_{282}							
30	4-Hydroxyazobenzene-2'-carboxylic acid	—	—		4.7	1.4		—	—	6×10^{-5}	7.2	—	7,8

Imidazolidone structure:

R_2
HN—1 2 3—NH
H 5 4 H
R_5 R_4

Oxazolidone structure:

O
O—1 2 3—NH
R_5 5 4 H
H R_4

[a] Note that C-3 and C-4 of biotin correspond to C-4 and C-5 of the imidazolidones. Since the latter were synthesized from α-amino acids, the α-CH of which became C-4 of the imidazolidone, they have been designated D or L according to their origin from the D or L series of amino acids. The absolute configuration of C-3 in biotin is the same as that of C-4 in the D-alkylimidazolidone. The relative configuration of the two oxazolidones is shown more clearly in Fig. 4. Further discussion of the stereochemistry and optical activity of some of these compounds is given by Green et al. (1970).

[b] The values of $\Delta E_{max}/E_{282}$ were determined with a Unicam SP. 700 spectrophotometer using avidin of specific activity 14. Redetermination of several of these values using pure avidin (specific activity 15) and a Cary 118 spectrophotometer gave values that were 10–15% higher (Table IV). The difference maxima were also shifted 1–2 nm to shorter wavelength, on account of the lower level of stray light from the double monochromator. The Unicam instrument gave wavelengths of maximum difference of 233 for all the compounds except 4, 18 (238 nm), 25, 27, 28, 29 (229 nm). The values of $\Delta E_{max}/E_{282}$ for urea and glycol have been corrected for a nonspecific red shift of the spectrum induced by the high concentration of reagent required to saturate the binding site. $\Delta E_{max}/E_{282}$ can be converted to a difference extinction coefficient per mole of analog bound by multiplying by 24,000.

[c] The values of k_{-1}/k_1 are based on the value of k_1 (7×10^7 M⁻¹sec⁻¹) determined for biotin at pH 5 (Green, 1963a). All measurements of rate and equilibrium constants were made at 25°C, except for those quoted from reference (2), which were at room temperature.

[d] The separate values of k_{-1} given for d and l isomers of dl desthiobiotin derivatives were calculated from the displacement rates of the dl mixtures, assuming in each case that the slower rate was that of the d isomer. The dl mixtures of the alkylimidazolidones gave linear Scatchard plots showing no significant difference in the affinity for d and l isomers.

[e] The value of K for the free base form was calculated from K measured at a lower pH, using the measured pK_a of the guanidinium group.

[f] ΔG and ΔH are the free energies and enthalpies of formation of the avidin complexes. They are all negative.

[g] 1, Green (1963a); 2, Green (1963b); 3, Green and Toms (1973); 4, Green and Toms (1975); 5, Suurkuusk and Wadsö (1972); 6, Green (1966); 7, Green (1965); 8, Green (1970).

stepwise modification of structure on the free energy of binding. The results obtained with close relatives of biotin confirmed the main conclusions of earlier workers concerning the relative importance of different parts of the biotin molecule. Biotin sulfone (compound 3), ester (compound 2) amides (Green, 1963b), and desthiobiotin (compound 8) all bind very firmly to avidin whereas the diamine (compound 23), in which the imidazolidone ring is broken, has a dissociation constant 10^7 times greater than that of biotin. Nevertheless, it was still bound quite firmly unless both amino groups were protonated. Other analogs (compounds 4 and 5) in which the imidazolone ring was less drastically modified were still bound strongly.

2. Stereochemistry of Biotin

Before considering the finer details of the effects of structure on binding, the stereochemistry of biotin and its derivatives will be outlined. The absolute configuration has been determined by X-ray crystallography (Trotter and Hamilton, 1966) and is shown in Table V. There are three asymmetric centers at C-2, C-3, and C-4. The two rings are fused in the cis configuration and the valeric acid side chain (C-2) is also cis in relation to the imidazolidone ring. There are four racemic pairs of stereoisomers (György, 1954). The binding of dl-epibiotin, epimeric at C-2, has not been studied. Wright and Skeggs (1947) showed that of the allo isomers (trans fusion of the rings) only the dl-epiallobiotin competed significantly with d-biotin (inhibition ratio, 6) and that l-biotin did not compete at all. The epiallo isomers resemble biotin in the relation of the valeric acid side chain at C-2 to N-1' but differ in its relation to N-3'.

In the series of imidazolidones and oxazolidones shown in Table V, only the asymmetric center at C-3 remains. This becomes C-4 in the imidazolidones and when its configuration is D (see footnote a to Table V) it resembles that of d-biotin. The methyl group of desthiobiotin produces an extra asymmetric center at C-5, with the opposite configuration to that at C-4 (imidazolidone numbering). If the two alkyl chains were equivalent, desthiobiotin would be an internally compensated meso compound.

3. Relation of Affinity to Structure

The results obtained with the 4-alkyl imidazolidones (compounds 11–16) and their analogs are summarized in a diagram (Fig. 4). The D-hexyl derivative, stereochemically similar to d-biotin, binds most firmly with ΔG (−11 kcal/mole) 2 kcal greater than that of the L-

D-Hexyl imidazolidone	H	CH₃	−11.1
L-Hexyl imidazolidone	H	CH₃	− 9.2
d-Desthiobiotin alcohol	CH₃	CH₂OH	−14
d-Desthiobiotin ester	CH₃	CO₂Et	−15.6
d-Desthiobiotin	CH₃	CO₂⁻	−17
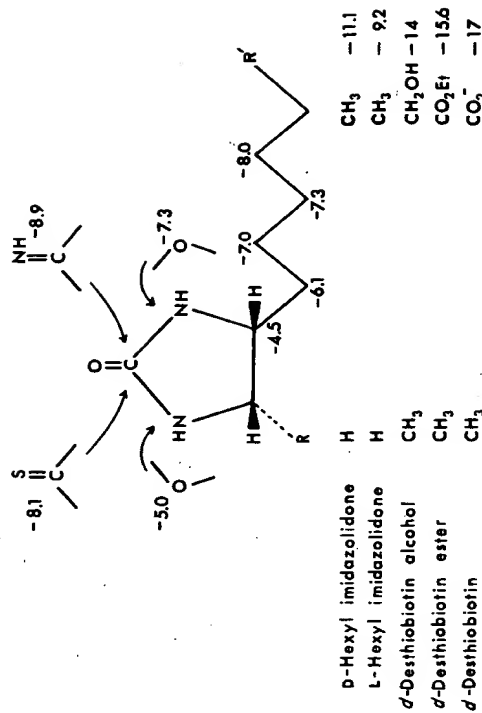
FIG. 4. Free energy of binding of hexyl imidazolidone and its analogs (Green and Toms, 1975). The numbers give the free energies of binding of a series of compounds regarded as derived from 4-n-hexylimidazolidone. The ring modifications were made one atom at a time using the D isomer. The series of shorter alkyl chains were all dl compounds.

isomer. When the alkyl chain was shortened, an approximately linear decrease in ΔG with decreasing chain length was observed implying that the whole of the alkyl chain was contributing to the binding energy. The average decrease of 0.7 kcal per methylene group was similar to that observed in hydrophobic interactions of other homologous series (Anderson et al., 1965; Smith and Tanford, 1973).

Much larger free-energy changes followed modification of the ring structure, the most dramatic being the drop of 6 kcal when N-1' was replaced by oxygen. This was considerably larger than that produced by a similar replacement at N-3'. A similar larger effect of modification of N-1' versus N-3' was seen when these positions in biotin were substituted with methoxycarbonyl groups (compounds 6 and 7). It seems likely that both NH groups are acting as H-bond donors to acceptor atoms on the protein and that the bond with N-1' is particularly strong. The small difference between D and L isomers (except for the hexyl derivative) can be understood in terms of the "meso" structure of desthiobiotin. Interchange of C-4 and C-5 by rotation about an axis through the carbonyl group would convert an analog with the "wrong" (L) configuration at C-4 to one with the right configuration at C-5. One might predict on this basis that the

L-isomer when bound to avidin would be rotated 180° with respect to the bound D-isomer. Introduction of a methyl group in the 5 position and a carboxyl group at the end of the alkyl chain gives desthiobiotin, further increasing the free energy of binding.

It appears that almost every atom in the biotin molecule contributes something to the free energy of interaction with avidin. This is confirmed by the ability of avidin to bind analogs corresponding to quite small fragments of biotin such as urea, glycol (equivalent to C-3, C-4, N-1', N-3'), tetrahydrofuran, and caproic acid. It is likely that each binds to a different part of the site since, for example, the spectral change produced by a mixture of imidazolidone and hexanoate is approximately the sum of the separate contributions (Green and Toms, 1975). It is also interesting that the free energy of binding of each of these two analogs was unaffected by the presence of an excess of the other. The sum of the free energies of binding of hexanoate (compound 29) and 4-methylimidazolidone (compound 15) is 11 kcal, considerably less than that of desthiobiotin (compound 8) (16.9 kcal), which is the equivalent single molecule. Thus although there is no evidence for cooperative interactions between different parts of the binding site, there is a considerable enhancement in the free energy of binding of biotin and its near analogs relative to that of the fragments.

This differs from the behavior of lactate dehydrogenase toward fragments of NAD (McPherson, 1970). The enzyme does not bind nicotinamide unless the other half of the coenzyme (AMP) is present, and under these conditions the sum of free energies of binding of the fragments (1.8 + 3.2 kcal) is close to that of the complete coenzyme (5.3 kcal).

Since binding of each fragment of biotin leads to red shifts of the tryptophan absorption spectrum and since each small modification of biotin decreases the red shift (Table V), it is likely that several tryptophans of each subunit interact directly with biotin. Further evidence for this comes from chemical modification experiments described in Section IV,C. Although there was a general trend toward a lower difference extinction coefficient as the free energy of binding decreased, the proportionality was not exact. For example, hydrolysis of the imidazolidone ring produced a large decrease in free energy of binding but only a small decrease in $\Delta\epsilon_{233}$ (compound 23), while removal of the sulfur from the thiophan ring had the converse effect (compounds 8, 9, 10). It is therefore difficult to draw any quantitative conclusions about the contribution of biotin–tryptophan interactions to the binding energy.

The wide range of compounds that will bind to avidin suggests a rather nonspecific binding site. But, apart from the dye (compound 30) and anilinonaphthalene sulfonate (Green, 1964b), they are all related to some part of the biotin molecule. On the other hand, a site of relatively high specificity is supported by studies on related compounds which did not bind significantly at the 1–10 mM level. These include the 20 natural amino acids, succinimide and imidazole-4-propionic acid (N. M. Green, unpublished experiments).

C. Periodate Oxidized Avidin

The effects of biotin on the spectrum of the formylkynureinine residue of oxidized avidin provided the opportunity of detecting differential effects of biotin analogs on at least three tryptophan residues. The difference spectrum produced by biotin (Fig. 5c) shows additional maxima at 242, 267, 276, and 345 nm, which are almost the same in position and relative magnitude as the peaks in the first derivative spectrum of α-N-acetyl-N-formylkynurenine, allowing for a red shift caused by the protein environment. Fortunately, they hardly overlap at all with the trypophan peaks so that differential effects are readily detectable. Examination of the spectra produced by analogs showed that esterification of the carboxyl group of biotin decreased the perturbation of the formylkynurenine residue without affecting the other tryptophans while removal of the sulfur atom showed the converse behavior. This implies that the oxidized tryptophan residue interacts in the neighborhood of the valeric acid side chain of biotin and that it is located away from the thiophan ring. One of the other tryptophans must be near the thiophan to account for the decreased $\Delta\epsilon_{233}$ when desthiobiotin was bound. The extent of the interaction of the valeric acid side chain with the formylkynurenine can be deduced from the different spectra given by propylimidazolidone, which produced a slight perturbation of the oxidized residue, and ethyl imidazolidone, which did not. Since ethyl imidazolidone gives a tryptophan difference spectrum it is likely that it interacts with a third tryptophan residue different from that which is perturbed by the thiophan ring.

D. Nature of the Binding Site

Although the strength of the avidin–biotin interaction suggests the possibility of a covalent attachment, most of the evidence favors noncovalent binding. The biotin can be released by 6 M GuHCl at pH 1.5 (Cuatrecasas and Wilchek, 1968) or by autoclaving. The study of analog binding shows extensive noncovalent interaction with all

that three or four tryptophan residues in each subunit interact directly with the biotin. In addition, there are groups which hydrogen bond to the ureido group and there is an amino group which reacts with FDNB to give a blocked binding site.

If this amino group were near to the tryptophan which is perturbed by the carboxyl group of biotin several other observations would fall into place. The susceptibility of this tryptophan to oxidation by periodate could be accounted for by binding of IO_4^- to NH_3^+. The observation that dinitrophenylation of this amino group prevents both the binding of biotin and the oxidation of the tryptophan by periodate (Green and Toms, 1975) would also be explicable. Although the amino group reacts readily with FDNB it does not react with the bulkier dansyl chloride and trinitrobenzene sulfonic acid reagents. This is consistent with a location near to the biotin carboxyl group, which studies with bisbiotinyl compounds (Section V,D) have shown to be some way beneath the van der Waals surface of the protein and hence subject to restricted access.

Since 70% of the biotin molecule is hydrophobic and much of it interacts with tryptophan, it was expected that the binding would be accompanied by an appreciable increase of entropy. However, calorimetric measurement of the enthalpy of binding (Green, 1966; Suurkuusk and Wadsö, 1972) led to the conclusion that there was no net entropic contribution to the free energy of binding. Nevertheless, the high negative ΔC_p (237 cal/°C/biotin) (Suurkuusk and Wadsö, 1972) confirmed a loss of hydrophobic surface accompanying binding. It was suggested that the high enthalpy of binding was a reflection of the formation of several strong hydrogen bonds to the imidazolidone ring. Since the entropy of formation of these bonds would be negative, the net entropy change might be very small. This interpretation is supported by the positive entropy of binding of the diamine (compound 23), in which the imidazolidone ring is broken. It would be possible to account for most of the large ΔH in these terms, only if the potential bonding atoms in avidin do not form hydrogen bonds to water when biotin is absent. It is difficult to see how this could happen without some local conformational change.

X-Ray crystallographic studies at present in progress could give a direct answer to this question. If the bound biotin analogs can be located with sufficient accuracy, it may also be possible to establish correlations between specific atomic interactions and the free energies of binding of analogs.

FIG. 5. Difference spectra of avidin or oxidized avidin complexes versus the ligand-free protein. The protein subunit concentration was 6.5 μM (0.4375 cm path) in 50 mM sodium phosphate pH 6.8 (Green, 1975). (a) Avidin-biotin; (b) first derivative of spectrum of α-N-acetyl-N-formylkynurenine; (c)-(g) Complexes of periodate oxidized avidin with a ligand, as follows: (c) biotin; (d) biotin methyl ester; (e) desthiobiotin; (f) propyl imidazolidone; (g) ethyl imidazolidone.

parts of the biotin molecule, and although the sum of the weak interactions is a significantly less than the free energy of binding of biotin the discrepancy is small enough to be explicable in terms of a slightly better fit of biotin as compared with the analogs. These results combined with those of chemical modification studies suggest

FIG. 6. Quenching of fluorescence of avidin by the DNP-hydrazide of biotin. The experimental points (Green, 1964b) are compared with theoretical curves recalculated for a four-subunit molecule from the relation $F/F_0 = \Sigma \, \alpha_i q_i$, where α_i is the fraction of avidin present as A_4, B_i and q_i is the relative quantum yield of the ith species. Knowing from the experimental points that $q_4 = 0.09$, it followed, from the fluorescence of avidin carrying covalently bound DNP (Green, 1964b), that $q_3 = 0.09$, $q_2 = 0.11$, $q_1 = 0.17$, $q_0 = 1$. Three curves were then calculated relating F/F_0 to x, the fractional saturation of the binding sites, using three different relationships between x and α_i: (1) sequential binding (left-hand line); (2) random binding (middle curve): α_i given by successive terms of the expansion of $[(1-x) + x]^4$; (3) cooperative binding (right-hand line): $\alpha_0 = (1-x)$, $\alpha_1 = \alpha_2 = \alpha_3 = 0$, $\alpha_4 = x$.

A similar analysis (Holbrook, 1972) of the quenching of the fluorescence of lactic dehydrogenase was consistent with random binding of NADPH by the four subunits. Here q_1/q_0 is considerably higher (0.59) than it is for the DNP-avidin system, since R_0 is smaller (25 Å) and the subunits are larger. In consequence the quenching curve is more nearly linear. By combination of appropriate curve-fitting procedures with accurate experimental results it was still possible to obtain information about the distribution of ligands among the binding sites.

V. SUBUNIT STRUCTURE

A. Dissociation in Guanidinium Chloride

Avidin resists unfolding in high concentrations of urea (Fraenkel-Conrat et al., 1952b) or moderate concentrations of GuHCl (Green,

E. Evidence for Random Binding

The strength of the interaction between avidin and biotin is so great that it would not have been surprising to find that binding of biotin at one site influenced the interaction at neighboring sites. Nevertheless, several recent observations have confirmed the early conclusion (Green, 1964a) based on quenching of fluorescence, that the binding is a random process and that there is no detectable interaction between the sites: (1) Scatchard plots of the binding of analogs ($K > 10^{-8}\,M$) are linear (Green and Toms, 1975). (2) The binding constants of L-4-hexylimidazolidone and of 2-hydroxy-azobenzene 4′-carboxylate were not altered when three quarters of the binding sites were presaturated with biotin (Green and Toms, 1975). (3) The heat of binding of biotin was not affected by partial presaturation of the binding sites (Green, 1966).

We will recapitulate the early evidence in more detail as it forms the basis of a general method for distinguishing cooperative from random binding that has recently been extended independently by Holbrook (1972) in a study of fluorescence quenching of dehydrogenases by NADH.

The method depends upon the ability of a dinitrophenylated biotin derivative to quench the fluorescence of subunits besides the one to which it is bound. This arises because R_0 [the distance between DNP and tryptophan at which quenching and emission have equal probability (30 Å)[1]] is comparable to the distance between binding sites (20–30 Å); even at a distance of 40 Å the probability of quenching is 16%. In consequence, the relation between fluorescence and fractional saturation of the sites will be nonlinear unless there is very strong positive or negative cooperativity that would eliminate species intermediate between A_4 and A_4 B_4. The shape of the quenching curve can be calculated for different models of the binding process provided that the quantum yields for the different intermediate species can be assigned. This can be done either by a curve-fitting procedure (Holbrook, 1972) assuming $q_i/q_0 = (q_1/q_0)^i$ or by a semiempirical procedure (Green, 1964b, and Fig. 6) in which the fluorescence of randomly labeled DNP-avidin was measured as a function of the number of covalently bound DNP groups. The good agreement between the experimental points and the theoretical relationship for random binding provides strong evidence against any cooperative interactions.

[1] The value of 35 Å given by Green (1964a) was based on the assumption of too high a quantum yield for tryptophan.

1963c). Above 3.5 M GuHCl the protein begins to unfold, as measured by loss of biotin binding activity, exposure of aromatic residues to the solvent and change of optical activity. At a concentration of 6 M GuHCl molecular weight measurement indicates dissociation to quarter-molecules. The rate of unfolding, measured by difference spectrophotometry, increased with the twelfth power of the concentration of GuHCl, and the kinetics could be accounted for as a sum of two exponential terms (recalculated from the published results). The loss of biotin-binding activity followed the same time course. These results could be accounted for in terms of a reversible unfolding of the protein with accumulation or reformation of a partially folded species, provided that the ultraviolet absorbance of the latter is the same as that of the fully unfolded form (Ikai et al., 1973). If the extinction coefficients of the different unfolded forms were not the same, the kinetics of loss of activity would differ from the kinetics of the spectral change. However, it seems unlikely that this can in fact be the explanation of the complex kinetics since, as we shall see, the unfolded species are not in equilibrium with native avidin.

When the transition and its reversal were followed as a function of GuHCl concentration by difference spectroscopy (Fig. 7), the regain of tertiary structure and biotin binding activity did not occur until the concentration of GuHCl was lowered to 2 M. The simplest explanation of apparent hysteresis in systems of this type is that insufficient time has been allowed for the system to come to equilibrium (Steinhardt and Zaiser, 1953; Tanford, 1968). This is not the explanation here since in 4.5 M GuHCl no further change was observed between 18 hours and 60 hours.

Further information on irreversibility came from the addition of biotin to the system at various stages of the cycle. Over the whole of the dissociation branch, the binding of biotin was immediate and no slow changes were observed, even when the system was left for 48 hours. The amount bound was proportional to the fraction of native avidin left as measured by the spectral change. If any other species had been present which could bind biotin, complete reversal of the reaction would have been expected, since monomeric avidin–biotin complex reassociates to native tetramer even in 6 M GuHCl (Green and Toms, 1972). This unexpected irreversibility was not the consequence of a slow secondary change, since the same result was obtained if biotin was added during the course of denaturation in 6 M GuHCl. The biotin reacted immediately with the remaining native protein, but did not react at all with the freshly formed unfolded protein. Dilution of the reaction mixture to a concentration of 3 M
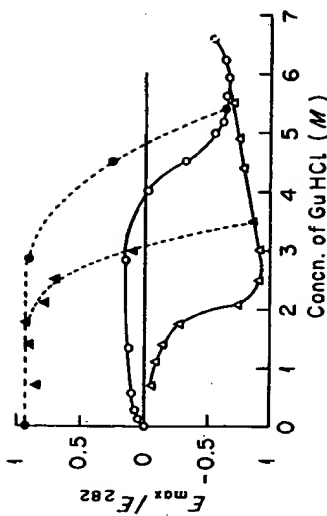
GuHCl, without addition of biotin, also gave a frozen system containing native tetramer and unfolded monomer.

In contrast, the reassociation branch of the curve behaved in a more normal manner. The ability to recombine with biotin was regained when the guanidine concentration reached 3 M; a slow binding was then observed. At this concentration, in the absence of biotin, molecular weight measurements (Green, 1964b) showed no detectable tetramer. This was consistent with later results (Green and Toms, 1973) which showed that single subunits of avidin could bind biotin. In this region of the curve there appears to be a reversible equilibrium between unfolded and partly refolded subunits which can be displaced when the latter interact with biotin. When the refolding progresses to the stage where tetramers can re-form the reaction goes rapidly to completion even in the absence of biotin, which could account for the very sharp transition in this region. The kinetics of refolding were first order and the rate was inversely proportional to the twenty-fifth power of the concentration of GuHCl. A second denaturation of the refolded avidin followed the same course as the first.

The behavior described is in many respects similar to that of systems which show genuine hysteresis, such as agar gels (Rees, 1969) or polyglutamic acid (Jennings et al., 1968) except that for these systems the temperature rather than the denaturant concentration was varied. It would be interesting to know whether avidin would show similar hysteresis with temperature at a constant GuHCl concentration in the region of 2 M.

The thermodynamics of these systems have been developed by Everett (Everett and Whitton, 1952; Everett, 1954, 1955), who pointed out that the transitions should be extremely sharp in homogeneous systems and that the finite width often observed is a consequence of the sequential melting of domains of different sizes. This is the one aspect of the behavior of avidin that is difficult to account for if we assume that it is a single molecular species. To elaborate this point, it seems most likely that at concentrations of GuHCl above 3 M the unfolded monomer represents the thermodynamically stable state. Nevertheless, the native tetramer is stable for periods of weeks in 3 M GuHCl and does not unfold at a significant rate until concentrations of 3.5–4 M are reached. It is difficult to understand why the unfolding does not go slowly to completion at this concentration, yet even at 4.5 M GuHCl it comes to a stop half way. It seems unlikely that the minor heterogeneity described by De Lange (1970) could account for the breadth of the transition, although it should be pointed out that heterogeneity could also explain the presence of two exponential terms in the unfolding kinetics. The refolding branch of the curve presents less of a problem since the transition is extremely sharp. Clearly more experimental work is required to resolve these contradictions, particularly on the effect of temperature on the system and on the formation of tertiary structure measured by CD.

Some time has been spent in this attempt to define the behavior of avidin in GuHCl since it is likely to be characteristic of the unfolding of other multisubunit proteins, few of which have been studied in detail. The refolding of aldolase after dissociation at low pH has been thoroughly examined by Teipel (1972). He showed a rapid regain of helical structure to give an inactive monomer–dimer mixture which slowly isomerized to a form which gave active tetramer. The second stage was accelerated by substrate. A similar sequence of events would be consistent with the more limited observations made with avidin, except that the inactive folded form of avidin is unlikely to be helical. The reversibility of the dissociation branch of the aldolase curve has not been studied, but preliminary results of Deal et al. (1963) using urea suggest that it may show irreversibility similar to that of avidin.

The dissociation of apoferritin at acid pH shows marked hysteresis, which has been followed by difference spectroscopy and by sedimentation velocity (Crichton and Bryce, 1973). Avidin also shows hysteresis at low pH (Green, 1963c).

B. Stabilization by Biotin

The effect of adding biotin during the process of unfolding has been considered in the preceding section. Although biotin could not induce refolding (Section III,A) even in 6 M GuHCl, it completely prevented unfolding (Section III,A) even in 8 M GuHCl or in 0.1 M HCl. A combination of 6 M GuHCl with low pH (1.5) does dissociate the complex and has been used to recover avidin from biocytinyl-Sepharose columns (Cuatrecasas and Wilchek, 1968).

1. Dissociation of Partially Saturated Avidin–Biotin Complexes

Further information about the pathway of unfolding of the tetrameric avidin and the nature of the stabilization by biotin has been obtained from measurements on avidin molecules partially saturated with biotin. Since subunits combine at random with biotin such preparations will be mixtures in which the number of occupied sites per molecule is binomially distributed. Occupied subunits present in a tetramer were resistant to 6 M GuHCl or to N-bromosuccinimide. The question was asked whether such subunits were equally resistant when neighboring subunits were unoccupied and, if so, whether they had any stabilizing influence on their neighbors. The answer to the first question was "yes" and to the second, "almost none," as will appear from the following section.

When avidin is treated with N-bromosuccinimide, oxidation of the tryptophan residues is followed by dissociation of the tetramer into subunits which partially aggregate (Green and Ross, 1968). The presence of biotin completely protected the subunit to which it was bound, but it did not affect the reactivity of neighboring subunits. The oxidized subunits dissociated, while the protected ones recombined to form tetrameric avidin–biotin complex. The effects of 6 M GuHCl were identical, except that the subunits were not oxidized (Green and Toms, 1972). The rate of unfolding of the vacant subunits in GuHCl was decreased by a factor of two in molecules which contained two or three occupied subunits, a small effect when considered in relation to the free energy of binding of biotin.

These observations have several further implications. Since all four tryptophan residues of the vacant subunits were oxidized and all those of the occupied subunits were protected, it follows that each binding site was almost certainly situated within a subunit rather than at an interface between subunits. This conclusion was also supported by the small effect of biotin bound to one subunit on the

stability of the neighboring subunits. This last observation also suggested that GuHCl initiated the unfolding in the region of the biotin binding site rather than at the intersubunit bonds. If the latter pathway were important, a greater effect of biotin on the unfolding of neighboring subunits would have been expected. Both sets of experiments imply that single subunits with bound biotin are stable, at least for short periods, at concentrations of GuHCl up to 6 M, since they must be intermediates in the overall reaction

$$A_4B \rightarrow 3A + AB$$
$$4AB \rightarrow A_4B_4$$
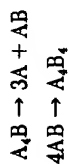
2. Thermal Stability

The thermodynamics of the dissociation of avidin has not yet been studied since the appropriate conditions for reversibility have not been established. Calorimetric measurements of the irreversible heat denaturation are of some interest, however, in relation to the stabilizing effect of biotin. By differential scanning calorimetry Donovan and Ross (1973) showed that avidin underwent a cooperative thermal transition at 85°C with a ΔH of 76 kcal per subunit. This is within the range normally found for denaturation of proteins. The corresponding transition in the avidin–biotin complex took place at 132°C under pressure and ΔH was increased, by almost 200 kcal, to 270 kcal per subunit. Some of this can be accounted for from the heat of binding of biotin, which could be as high as 46 kcal per subunit at 132°C (ΔC_p is 240 cal/°C per subunit, Suurkuusk and Wadsö, 1972). Donovan and Ross pointed out that most of the remaining difference could also be an effect of the high transition temperature and the large measured ΔC_p of 3.0 kcal/°C per subunit.

C. Binding of Biotin by Single Subunits

Single subunits have been obtained free in solution only in the presence of GuHCl or at acid pH. To study their properties under normal conditions it was necessary to have them coupled to a Sepharose matrix to prevent them from reassociating (Green and Toms, 1973). Low concentrations of tetramer were coupled using (1) small amounts of cyanogen bromide, to minimize the proportion of doubly linked molecules and (2) low concentrations of tetramer, to avoid interactions between the covalently coupled subunits after the 75% of noncovalently bound subunits had been removed with 6 M GuHCl. In spite of these precautions the flexibility of the Sepharose matrix permitted some interaction between covalently linked subunits. In

consequence the preparation showed three classes of binding site in approximately equal proportions. The strongest was indistinguishable from tetrameric avidin while the weakest ($K = 10^{-7} M$) was assigned to single subunits. The proportion of the latter was doubled when the Sepharose was stabilized by cross-linking with divinylsulfone. All classes of covalently bound subunits were able to reform tetramers when subunits in 3 M GuHCl were added and the system was diluted. It can be concluded that the Sepharose matrix did not interfere with refolding of the peptide chain and that the formation of active subunits was not dependent on interactions with other subunits. Subunit interaction, however, was required for firm binding of biotin.

The question of whether there exists a finite concentration of single subunits in equilibrium with the tetramer has not been answered. The most sensitive method for detecting such an equilibrium is to look for formation of hybrids after long incubation. Preliminary experiments in the author's laboratory have used mixtures of trinitrophenyl avidin (TNP$_4$A$_4$) and unlabeled avidin. The fluorescence of the mixture should decrease by half when a normal subunit is exchanged for a TNP subunit. Less than 5% change in fluorescence was observed over periods of several days showing that the dissociation constant for formation of monomers or dimers from tetramers must be less than $10^{-12} M$. [If a rate constant of $10^6 M^{-1}$ sec^{-1} is assumed for the association reaction between macromolecular species an approximate dissociation constant (K) can be calculated from the rate of subunit exchange (= rate of dissociation), $K = 10^{-6} k$.]

Further experiments along these lines in the presence of GuHCl could provide information on the species which are in equilibrium with each other in the different states illustrated in Fig. 7.

D. Bifunctional Biotin Derivatives and Subunit Structure

Although examination of avidin in the electron microscope did not reveal any subunits (Fig. 8A), it was possible to deduce the symmetry of their arrangement from the morphology of polymers produced when avidin combined with bisbiotinyldiamines (Green et al., 1971). These polymers were formed only when the chain joining the carboxyl groups of biotin was more than 15 Å long. When the chain length was increased to 18 Å the second biotin was bound essentially irreversibly. With intermediate chain lengths it was possible to depolymerize the products with an excess of reagent or with the dye 4-hydroxyazobenzene 2'-carboxylate. The polymers (Fig.

8b) were linear, unbranched, and as wide as a single avidin molecule (55 Å). The repeat distance of 41 Å suggested that each molecule was a cylinder or an oblate ellipsoid 55 Å × 41 Å, consistent with the size of the unit cell of the crystal (62 Å × 107 Å × 43 Å) which contained two molecules (Green and Joynson, 1970). A tetramer with 2-fold symmetry would give linear polymers when linked by chains of minimal length, whereas one with 4-fold symmetry would not. Since spectrophotometric titration showed that all the sites were saturated, each molecule was doubly linked to its neighbor and the binding sites were arranged in pairs on the 55 Å faces as indicated in the diagram (Fig. 9). When linking chains longer than 23 Å were used the length of the polymers was markedly decreased, suggesting termination by intramolecular binding of the bifunctional ligand.

In the stable polymers the carboxyl group of the bound biotin lies 8–9 Å beneath the van der Waals surface of the molecule. Even so it must be accessible to the solvent since large molecules can be coupled to it without disturbing the binding. This suggests that there may be a shallow depression in the surface as illustrated in Fig. 9. This would account for the ability of a 23 Å chain to bridge neighboring sites without bringing the sites improbably close together.

When avidin was combined with four molecules of 1-biotinamido-12-dinitrophenylaminododecane only two of the four DNP groups were titratable with Fab fragments of anti-DNP antibody and with few exceptions only two Fab fragments could be seen bound to each avidin in electron micrographs of the product (Fig. 8C) (Green, 1972). This confirms the 2-fold symmetry of the avidin molecule and shows that the members of each pair of sites are too close together to accommodate 2 Fab fragments (diameter 40 Å) without considerable strain. The occasional avidin carrying three Fab fragments suggests that weak binding of extra Fabs is possible.

An extension of this experimental approach which may prove useful is to employ avidin as a bifunctional reagent for linking biotin labeled macromolecules together. The biotin could be linked directly to a lysine residue of a protein or combined with a firmly bound specific ligand. The products may be useful for electron

FIG. 8. Electron micrographs of avidin. (A) Avidin; (B) avidin polymerized with bisbiotinamidododecane; (C) avidin labeled with Fab fragments of anti-DNP antibody using 1-dinitrophenylamino-12-biotinamidododecane. The molecules with three Fab fragments (arrow) represented 2 or 3% of the total population. ×425,000. From Green (1972).

radioactive biotin or on spectral changes that accompany biotin binding.

Since avidin is strongly adsorbed by carboxymethyl cellulose (Green, 1963a) or by bentonite (Fraenkel-Conrat et al., 1952a), the amount of bound radioactive biotin is easily determined. The adsorption by bentonite is more efficient and the sensitivity of the method is limited only by the specific activity of the biotin (Korenman and O'Malley, 1967) and can be of the same order as that of the microbiological assay. The estimation of radioactive avidin–biotin complex by gel filtration (Wei and Wright, 1964) should be used with caution since avidin is adsorbed by Sephadex, particularly at low salt concentrations. The effect should be small in 0.2 M ammonium carbonate, as used by Wei and Wright, but it can be significant in 0.1 M NaCl at pH 7 (N. M. Green, unpublished observations).

The spectrophotometric method is more convenient but much less sensitive (20 μg of avidin per milliliter, 10^{-6} M). It is based on the large spectral change in the dye 4-hydroxyazobenzene-2'-carboxylic acid when it is bound to avidin (Green, 1965). A new absorption band appears ($\epsilon_{500} = 34,000$) characteristic of the quinonoid form of the dye, and this disappears when the dye is displaced by biotin. The amount of avidin may be calculated directly from the increased absorbance at 500 nm or the dye may be used as an indicator in a spectrophotometric titration with biotin. Since the extinction coefficient and the dissociation constant of the avidin–dye complex remain constant between pH 4 and pH 8 and over a wide range of salt concentrations, the method is very convenient (Green, 1970). It can be used to measure protein bound biotin, preferably after partial digestion with pronase. Satisfactory results have been obtained both with purified proteins (Scrutton and Mildvan, 1968; Gerwin et al., 1969) and with chloroplast preparations (Kannangara and Stumpf, 1973). The only misleading system (Wellner et al., 1968; Huston and Cohen, 1969) also gave misleading results when microbiological procedures were used. Further details of specificity and experimental procedure for this and other methods of assay for avidin and biotin can be found elsewhere (McCormick and Wright, 1970).

B. Avidin and Biotinyl Enzymes

Avidin has been used in the study of several different aspects of biotinyl enzymes. In the first place it has provided a criterion for the involvement of biotin in enzymic reactions. This is particularly useful since biotin is always covalently bound and is difficult to detect by other means. In detailed studies of the molecular character-

FIG. 9. Structure of avidin polymers with bisbiotinyldiamines. The location of the biotin carboxyls (●) relative to the surface is defined by the minimal chain length of diamine required for stable polymers (18 Å). Their intramolecular separation is defined by the chain length (23 Å) leading to termination of the polymers and by the depth of the depression in which the sites are situated. (A) Polymer chain viewed end on. The sites on the lower face of the terminal molecule are shown as open circles. The rotation of one molecule relative to the next is determined by the angle between pairs of sites. In the alternative side views, (B) and (C), this angle has been arbitrarily set at 0° to simplify illustration. The 20 Å separation of sites in (B) requires a deeper depression than the 15 Å separation in (C). Arrows and elongated solid ellipses represent 2-fold symmetry axes. From Green et al. (1971).

microscopy, for studies of energy transfer in fluorescent systems or for immunological studies of linked antigens. Provided that the macromolecule is more than 40 Å in diameter, no more than one should combine on each face of the avidin molecule.

VI. MISCELLANEOUS TOPICS

A. Methods for Estimation of Avidin

The early methods for estimation of avidin were based on microbiological assay of biotin using mainly yeast (Hertz, 1943) or L. arabinosus (Wright and Skeggs, 1944) as the test organism. They are discussed in detail elsewhere (Snell, 1950; György, 1954). They are extremely sensitive, being capable of measuring biotin concentrations of 10^{-11} M, but are rather time consuming. A number of more convenient methods are now available based either on the use of

Enzyme	Ligand added	Approximate rate constant[a] $(M^{-1} \text{sec}^{-1})$	Reference
Methylmalonyl-CoA pyruvate transcarboxylase	—	10^6	N. M. Green (unpublished)
Propionyl-CoA carboxylase	—	5×10^4 [b]	Kaziro et al. (1960)
Acetyl-CoA carboxylase	—	10^4 to 10^5	Moss and Lane (1972)
Pyruvate carboxylase	Citrate	0	Scrutton and Utter (1965)
	—	500	Scrutton and Utter (1965)
	ATP	0	
	Acetyl-CoA	2500	Scrutton and Utter (1967)

TABLE VI

Rate of Inhibition of Biotinyl Enzymes by Avidin

[a] Except for propionyl-CoA carboxylase the second-order rate constants were calculated from first-order rate constants measured in the presence of excess avidin.

[b] The rate of inhibition was measured in the presence of one equivalent of avidin. A previously published figure (Green, 1963a) calculated from the same results was overestimated by a factor of 60.

istics of biotinyl enzymes avidin has been used to measure changes in the availability of biotin in the presence of various ligands and conversely to measure the binding of these ligands by the enzyme. It has also been used to locate the biotinyl residue in electron micrographs of complex enzymes (Green et al., 1972). It may yet prove possible to employ it or some modified version of lower affinity for enzyme purification.

The biotinyl enzymes themselves have been comprehensively reviewed by Moss and Lane (1971), and only those points specifically related to avidin will be considered here. Before 1950, studies of biotin-deficient organisms had implicated biotin more or less directly in a number of metabolic pathways. Some of the first direct evidence for its involvement in a specific enzyme came from a study of the effect of avidin in blocking incorporation of CO_2 into oxaloacetate (Wessman and Werkman, 1950). However, it was not until 1960 that the enzyme responsible, pyruvate carboxylase, was purified and shown to contain biotin (Utter and Keech, 1960). Meanwhile two other biotinyl enzymes which carboxylated acetyl-CoA and β-methylcrotonyl-CoA had been characterized (Wakil et al., 1958; Lynen et al., 1959) and had been shown to be inhibited by avidin.

1. Kinetics of Inhibition

Since avidin has four binding sites and most biotinyl enzymes possess at least four subunits, complex kinetics changing with molar ratio of the reactants are to be expected. Although the limited studies that have been made confirm this expectation, it has still been possible to draw some interesting conclusions from the results. When only a small excess of pure avidin was used an initial fast stage of inhibition followed by a slow second stage was observed for both pyruvate carboxylase (Scrutton and Mildvan, 1968) and transcarboxylase (N. M. Green, unpublished experiments). With a larger excess of avidin the inhibition of pyruvate carboxylase was first order in time but of order 1.5 with respect to avidin concentration (Scrutton and Utter, 1965). These effects probably reflect the steric barriers to reactions within and between polymeric complexes. Approximate values of rate constants (Table VI) show considerable differences between different enzymes and some interesting effects of ligands on the availability of biotin to avidin. When acetyl-CoA carboxylase is activated by isocitrate it forms polymers that do not react with avidin (Ryder et al., 1967). Conversely, avidin prevents activation and polymer formation if added before isocitrate. Since it does not reverse the polymerization, it can be used to follow the

kinetics of polymerization and depolymerization (Moss and Lane, 1972).

The availability of the biotin of pyruvate carboxylase to avidin is less than that of biotin in other enzymes, but it is increased by the allosteric activator, acetyl-CoA. Other ligands (oxalate, pyruvate, oxaloacetate) also have small effects on the rate which have been used to determine their dissociation constants (Mildvan et al., 1966).

The resistance of the activated form of acetyl-CoA carboxylase to inhibition by avidin provides the only example of such behavior of a biotinyl enzyme. It shows that resistance to avidin does not by itself eliminate the possibility that an enzyme reaction may involve biotin. There was one apparent example of the converse behavior. The synthesis of δ-aminolevulinate involved a decarboxylation step and the synthetase enzyme was inhibited by impure avidin. However, the inhibition was not prevented by biotin, and more highly purified avidin had no effect (Gibson et al., 1962).

2. Reversibility of Inhibition

In all biotinyl enzymes the biotin is covalently linked to a protein lysine and in almost all it acts as a carrier of CO_2 between substrates on two different active sites each catalyzing half of the overall reac-

tion. The biotin has to be free to move through appreciable distances (at least 10–20 Å), so it is not surprising that it is available to avidin and that the inhibition is usually irreversible. [If "irreversible" means less than 10% reactivation in 1 hour, this implies a dissociation rate constant of $< 3 \times 10^{-5}$ sec^{-1} and a dissociation constant of $< 3 \times 10^{-11}$ M. (cf. Section V,C).]

There are some reports of a partial reversal of avidin inhibition (Halenz et al., 1962; Scrutton and Mildvan, 1968) and even of complete reversal of one of the half-reactions catalyzed by propionyl-CoA carboxylase (Halenz et al., 1962; Friedman and Stern, 1961). Restoration of a half-reaction could be an effect of noncovalently bound biotin which might bind and be carboxylated in the absence of competition from the covalently bound biotin, now held by the avidin. Such a mechanism would be consistent with the high concentration of biotin (1 mM) required to restore activity. The partial reactivation of pyruvate carboxylase has been studied in some detail (Scrutton and Mildvan, 1968) and has a different origin. It was observed only when minimal concentrations of avidin were used to inhibit the enzyme. Under these conditions the reaction products contained a mixture of species of high molecular weight, in which the tetrameric enzyme was linked together by avidin molecules. The system could be depolymerized by an excess of avidin or by biotin. The latter restored up to 50% of the original activity while excess avidin rendered the inhibition irreversible. It appeared that when avidin was involved in cross-bridging of enzyme molecules it was unable to bind firmly, a situation reminiscent of the behavior of avidin polymers made with the shorter chain bisbiotinyl compounds (Section V,D). These results confirm that the biotin residues of pyruvate carboxylase are less accessible to avidin than they are in most other enzymes.

3. Carboxylated Enzymes

Avidin is able to combine with N-carboxylated derivatives of biotin (Table V) with a reduced affinity. It also inactivates the carboxylated form of propionyl-CoA carboxylase (Kaziro and Ochoa, 1961), but no extensive experiments with other enzymes have been reported. The inhibition may in part arise from the avidin increasing the rate of decarboxylation of the carboxybiotinyl enzyme. Edwards and Lane (quoted by Moss and Lane, 1971) have found that the rate of decarboxylation of acetyl-CoA carboxylase is increased several hundred times by avidin. It is likely that this enhanced rate is a consequence of a higher pK of the carboxyl group in the hydro-

phobic environment of the biotin-binding site. The protonated form of N-carboxyimidazolidone decarboxylates several thousand times faster than the anion (Caplow and Yager, 1967), so an increased pK would increase the rate of decarboxylation at a given pH.

4. Isolation of Biotinyl Enzymes

The use of avidin affinity columns to purify biotinyl enzymes remains an unfulfilled dream; no method for releasing the enzyme in a native state has been found. The combination of 6 M GuHCl and pH 1.5 has been used to release biotinyl peptides from avidin-Sepharose columns, but the recovery was only 24% (Bodanszky and Bodanszky, 1970). The biotin-containing subunit of acetyl-CoA carboxylase has also been obtained in unspecified yield by this method (Moss and Lane, 1971; Landman and Dakshinamurti, 1973).

Several attempts have been made in the author's laboratory to use either avidin subunits (Green and Toms, 1973) or periodate-oxidized avidin (N. M. Green, unpublished) coupled to Sepharose 4B to isolate biotinyl enzymes. Although the affinity of these derivatives for biotin was lower than that of avidin ($K = 10^{-7}$ to 10^{-9} M), it was still too high to allow significant release of enzyme, possibly because of multiple binding of several biotin residues on each molecule. There is still hope that further selective modification of the binding site may give derivatives with sufficiently low affinity to be useful for enzyme purification.

It is worth noting that avidin can be used as a specific adsorbent without covalent coupling since its high isoelectric point leads to strong adsorption on CM-cellulose (Green, 1963a).

The significant affinity of avidin for lipoate (Table V) suggests that avidin columns might be of use in the isolation of lipoyl enzymes or peptides.

C. Biosynthesis of Avidin

Certain aspects of avidin biosynthesis have attracted much attention recently. It had long been known that avidin synthesis was specifically stimulated by progesterone (Hertz et al., 1943) and this led O'Malley and his colleagues to a detailed study summarized by O'Malley et al. (1969). The main conclusions were that avidin is synthesized in the goblet cells of the epithelium of the oviduct whereas the other major proteins of egg white are synthesized in the underlying tubular gland cells and that this accounts for the differential effects of estrogen and progesterone on the two classes of protein (Kohler et al., 1968). Progesterone probably acts by control of tran-

scription of the avidin mRNA and recent work has been concentrated on this aspect of the process. The mRNA has been partially purified (Rosenfeld et al., 1972) and its synthesis in response to progesterone has been followed (Chan et al., 1973). Introduction of the partially purified mRNA into the estrogen primed oviduct stimulated synthesis of avidin in the absence of progesterone (Tuohimaa et al., 1972).

D. Streptavidin

No detailed characterization of this interesting relative of avidin has been published since the papers describing its discovery and isolation (Chaiet et al., 1963; Chaiet and Wolf, 1964; Tausig and Wolf, 1964). In spite of considerable differences in composition (Table 1) it is in most other respects remarkably similar to avidin from eggs.

It has a high tryptophan content, which gives rise to positive bands in the CD spectrum at 280 and 230 nm (Green and Melamed, 1966) that are characteristic of avidin and hitherto unique to these proteins. The tryptophan residues are less susceptible to oxidation by N-bromosuccinimide than are those of avidin. Biotin produces the same characteristic different spectrum when it binds and it protects the tryptophan against oxidation by N-bromosuccinimide. One or two of the tryptophan residues are susceptible to oxidation by periodate and biotin binds to the product giving difference peaks characteristic of formylkynurenine (N. M. Green, unpublished experiments). The heat of binding of biotin (23 kcal/mole) is similar to that shown by avidin (Green, 1966).

The molecular weight of streptavidin is about 60,000 and 1 mole of biotin is bound per 14,000 g (17-18 µg/mg streptavidin, Chaiet and Wolf, 1964). It is even more resistant than avidin to dissociation into subunits by guanidinium chloride.

No detailed study has been made of its binding of biotin analogs, but Lichstein and Birnbaum (1965) used microbiological methods to show that substitution in the carboxyl group does not affect the binding, but that neither the diamine (compound 23) nor desthiobiotin compete significantly with biotin, so that both rings must be intact for strong binding. Like avidin, it gives a complex with four moles of the dye 2-hydroxyazobenzene 4'-carboxylate ($K = 10^{-4}\ M$), which has the same extinction coefficient as the avidin complex (Green, 1970).

REFERENCES

Alderton, G., Lewis, J. C., and Fevold, H. L. (1945). Science 101, 151.
Ananthanarayanan, V. S., and Bigelow, C. C. (1969). Biochemistry 8, 3717.
Anderson, B. M., Reynolds, M. L., and Anderson, C. D. (1965). Biochim. Biophys. Acta 99, 46.
Andrews, L. J., and Forster, L. S. (1972). Biochemistry 11, 1875.
Baggaley, K. H., Blessington, B., Falshaw, C. P., Ollis, W. D., Chaiet, L., and Wolf, F. J. (1969). Chem. Commun. 3, 101.
Boas, M. (1927). Biochem. J. 21, 712.
Bodanszky, A., and Bodanszky, M. (1970). Experientia 26, 327.
Canfield, R. (1963). J. Biol. Chem. 238, 2698.
Caplow, M., and Yager, M. (1967). J. Amer. Chem. Soc. 89, 4513.
Chaiet, L., and Wolf, F. J. (1964). Arch. Biochem. Biophys. 106, 1.
Chaiet, L., Miller, R. W., Tausig, F., and Wolf, F. J. (1963). Antimicrob. Ag. Chemother. 3, 28.
Chan, L., Means, A. R., and O'Malley, B. W. (1973). Proc. Nat. Acad. Sci. U. S. 70, 1870.
Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972). Biochemistry 11, 4120.
Chou, P. Y., and Fasman, G. D. (1974). Biochemistry 13, 222.
Crichton, R. R., and Bryce, C. F. H. (1973). Biochem. J. 133, 289.
Cuatrecasas, P., and Wilchek, M. (1968). Biochem. Biophys. Res. Commun. 33, 235.
Dayhoff, M. O. (1972). In "Atlas of Protein Sequence and Structure," Vol. 5, p. 109. Nat. Biomed. Res. Found., Washington, D. C.
Deal, W. C., Rutter, W. J., and Van Holde, K. E. (1963). Biochemistry 2, 246.
De Lange, R. J. (1970). J. Biol. Chem. 245, 907.
De Lange, R. J., and Huang, T.-S. (1971). J. Biol. Chem. 246, 698.
Dhyse, F. G. (1954). Proc. Soc. Exp. Biol. Med. 85, 515.
Dittmer, K., and du Vigneaud, V. (1944). Science 100, 129.
Dittmer, K., du Vigneaud, V., György, P., and Rose, C. S. (1944). Arch. Biochem. 4, 229.
Dolby, L. J., and Booth, D. C. (1966). J. Amer. Chem. Soc. 88, 1049.
Donovan, J. W., and Ross, K. D. (1973). Biochemistry 12, 512.
Eakin, R. E., Snell, E. E., and Williams, R. J. (1941). J. Biol. Chem. 140, 535.
Everett, D. H. (1954). Trans. Faraday Soc. 50, 1077.
Everett, D. H. (1955). Trans. Faraday Soc. 51, 1551.
Everett, D. H., and Whitton, W. I. (1952). Trans. Faraday Soc. 48, 749.
Fraenkel-Conrat, H., Snell, N. S., and Ducay, E. D. (1952a). Arch. Biochem. Biophys. 39, 80.
Fraenkel-Conrat, H., Snell, N. S., and Ducay, E. D. (1952b). Arch. Biochem. Biophys. 39, 97.
Friedman, D. L., and Stern, J. R. (1961). Biochem. Biophys. Res. Commun. 4, 266.
Gerwin, B. I., Jacobson, B. E., and Wood, H. G. (1969). Proc. Nat. Acad. Sci. U. S. 64, 1315.
Gibson, K. D., Neuberger, A., and Tait, G. H. (1962). Biochem. J. 83, 539.
Gould, H. J., Gill, T. J., and Doty, P. (1964). J. Biol. Chem. 239, 2842.
Green, N. M. (1962). Biochim. Biophys. Acta 59, 244.
Green, N. M. (1963a). Biochem. J. 89, 585.
Green, N. M. (1963b). Biochem. J. 89, 599.
Green, N. M. (1963c). Biochem. J. 89, 609.
Green, N. M. (1964a). Biochem. J. 90, 564.
Green, N. M. (1964b). Biochem. J. 92, 16C.
Green, N. M. (1965). Biochem. J. 94, 23C.
Green, N. M. (1966). Biochem. J. 101, 774.
Green, N. M. (1968). Nature (London) 217, 254.

McPherson, A. (1970). *J. Mol. Biol.* **51**, 39.

Melamed, M. D., and Green, N. M. (1963). *Biochem. J.* **89**, 591.

Mildvan, A. S., Scrutton, M. C., and Utter, M. F. (1966). *J. Biol. Chem.* **241**, 3488.

Miller, A. K., and Tausig, F. (1964). *Biochem. Biophys. Res. Commun.* **14**, 210.

Montgomery, R. (1972). *In* "Glycoproteins" (A. Gottschalk, ed.), BBA Library, Vol. 5A, p. 519. Elsevier, Amsterdam.

Moss, J., and Lane, M. D. (1971). *Advan. Enzymol.* **35**, 321.

Moss, J., and Lane, M. D. (1972). *J. Biol. Chem.* **247**, 4944.

Mudd, J. B., Leavitt, R., Ongun, A., and McManus, T. T. (1969). *Atmos. Environ.* **3**, 669.

O'Malley, B. W., McGuire, W. L., Kohler, P. O., and Korenman, S. G. (1969). *Recent Prog. Horm. Res.* **25**, 105.

Pai, C. H., and Lichstein, H. C. (1964). *Proc. Soc. Exp. Biol. Med.* **116**, 197.

Pennington, D., Snell, E. E., and Eakin, R. E. (1942). *J. Amer. Chem. Soc.* **64**, 469.

Pritchard, A. B., McCormick, D. B., and Wright, L. D. (1966). *Biochem. Biophys. Res. Commun.* **25**, 524.

Rees, D. A. (1969). *Advan. Carbohyd. Chem. Biochem.* **24**, 267.

Rosenfeld, G. C., Comstock, J. P., Means, A. R., and O'Malley, B. W. (1972). *Biochem. Biophys. Res. Commun.* **47**, 387.

Rosenheck, K., and Doty, P. (1961). *Proc. Nat. Acad. Sci. U. S.* **47**, 1775.

Ryder, E., Gregolin, C., Chang, H., and Lane, M. D. (1967). *Proc. Nat. Acad. Sci. U. S.* **57**, 1455.

Scrutton, M. C., and Mildvan, A. S. (1968). *Biochemistry* **7**, 1490.

Scrutton, M. C., and Utter, M. F. (1965). *J. Biol. Chem.* **240**, 3714.

Scrutton, M. C., and Utter, M. F. (1967). *J. Biol. Chem.* **242**, 1723.

Smith, R., and Tanford, C. (1973). *Proc. Nat. Acad. Sci. U. S.* **70**, 289.

Snell, E. E. (1950). *In* "Vitamin Methods" (P. György, ed.), Vol. 1, p. 327. Academic Press, New York.

Steinhardt, J., and Zaiser, E. M. (1953). *J. Amer. Chem. Soc.* **75**, 1599.

Strickland, E. H., Horwitz, J., and Billups, C. (1969). *Biochemistry* **8**, 3205.

Suurkuusk, J., and Wadsö, I. (1972). *Eur. J. Biochem.* **28**, 438.

Tanford, C. (1968). *Advan. Protein Chem.* **23**, 218.

Tanford, C. (1970). *Advan. Protein Chem.* **24**, 1.

Tausig, F., and Wolf, F. J. (1964). *Biochem. Biophys. Res. Commun.* **14**, 205.

Teale, F. W. J. (1960). *Biochem. J.* **76**, 381.

Teipel, J. W. (1972). *Biochemistry* **11**, 4100.

Trotter, J., and Hamilton, J. A. (1966). *Biochemistry* **5**, 713.

Tuohimaa, P., Segal, S. J., and Koide, S. (1972). *Proc. Nat. Acad. Sci. U. S.* **69**, 2814.

Utter, M. F., and Keech, D. B. (1960). *J. Biol. Chem.* **235**, PC17.

Wahl, P., and Weber, G. (1967). *J. Mol. Biol.* **30**, 371.

Wakil, S. J., Titchener, E. B., and Gibson, D. M. (1958). *Biochim. Biophys. Acta* **29**, 225.

Wei, R. D., and Wright, L. D. (1964). *Proc. Soc. Exp. Biol. Med.* **117**, 341.

Wei, R. D., Kou, D. H., and Hoo, S. L. (1971). *Experientia* **27**, 366.

Wellner, V. P., Santos, J. I., and Meister, A. (1968). *Biochemistry* **7**, 2848.

Wessman, G. E., and Werkman, C. H. (1950). *Arch. Biochem.* **26**, 214.

Woolley, D. W., and Longsworth, L. G. (1942). *J. Biol. Chem.* **142**, 285.

Wright, L. D., and Skeggs, H. R. (1944). *Proc. Soc. Exp. Biol. Med.* **56**, 95.

Wright, L. D., and Skeggs, H. R. (1947). *Arch. Biochem.* **12**, 27.

Wright, L. D., Skeggs, H. R., and Cresson, E. L. (1951). *J. Amer. Chem. Soc.* **73**, 4144.

Green, N. M. (1970). *In* "Methods in Enzymology" (D. B. McCormick and L. D. Wright, eds.), Vol. 18A, p. 418. Academic Press, New York.

Green, N. M. (1972). *In* "Mosbach Colloquium" (R. Jaenicke and E. Helmreich, eds.), No. 23, p. 183. Springer-Verlag, Berlin and New York.

Green, N. M. (1975). In preparation.

Green, N. M., and Joynson, M. A. (1970). *Biochem. J.* **118**, 71.

Green, N. M., and Melamed, M. D. (1966). *Biochem. J.* **100**, 614.

Green, N. M., and Ross, M. E. (1968). *Biochem. J.* **110**, 59.

Green, N. M., and Toms, E. J. (1970). *Biochem. J.* **118**, 67.

Green, N. M., and Toms, E. J. (1972). *Biochem. J.* **130**, 707.

Green, N. M., and Toms, E. J. (1973). *Biochem. J.* **133**, 687.

Green, N. M., and Toms, E. J. (1975). In preparation.

Green, N. M., Mose, W. P., and Scopes, P. M. (1970). *J. Chem. Soc., C* pp. 1330–1333.

Green, N. M., Konieczny, L., Toms, E. J., and Valentine, R. C. (1971). *Biochem. J.* **125**, 781.

Green, N. M., Valentine, R. C., Wrigley, N. G., Ahmad, F., Jacobson, B. E., and Wood, H. G. (1972). *J. Biol. Chem.* **247**, 6284.

György, P. (1954). *In* "The Vitamins" (W. H. Sebrell, Jr. and R. S. Harris, eds.), Vol. 1, p. 527. Academic Press, New York.

György, P., and Rose, C. S. (1941). *Science* **94**, 261.

György, P., Rose, C. S., Hofmann, K., Melville, D. B., and du Vigneaud, V. (1940). *Science* **92**, 609.

György, P., Rose, C. S., and Tomarelli, R. (1942). *J. Biol. Chem.* **144**, 169.

Halenz, D. R., Feng, J., Hegre, C. S., and Lane, M. D. (1962). *J. Biol. Chem.* **237**, 2140.

Hammes, G. G. (1969). *Advan. Protein Chem.* **23**, 1.

Hertz, R. (1943). *Proc. Soc. Exp. Biol. Med.* **52**, 15.

Hertz, R. (1946). *Physiol. Rev.* **26**, 479.

Hertz, R., Fraps, R. M., and Sebrell, W. H. (1943). *Proc. Soc. Exp. Biol. Med.* **52**, 142.

Holbrook, J. J. (1972). *Biochem. J.* **128**, 921.

Huang, T.-S. (1971). Ph. D. Thesis, University of California, Los Angeles.

Huang, T.-S., and De Lange, R. J. (1971). *J. Biol. Chem.* **246**, 686.

Huston, R. B., and Cohen, P. P. (1969). *Biochemistry* **8**, 2658.

Ikai, A., Fish, W. W., and Tanford, C. (1973). *J. Mol. Biol.* **73**, 165.

Jennings, B. R., Spach, G., and Schuster, T. M. (1968). *Biopolymers* **6**, 635.

Kannangara, C. G., and Stumpf, P. K. (1973). *Arch. Biochem. Biophys.* **155**, 391.

Kaziro, Y., and Ochoa, S. (1961). *J. Biol. Chem.* **236**, 3131.

Kaziro, Y., Leone, E., and Ochoa, S. (1960). *Proc. Nat. Acad. Sci. U. S.* **46**, 1319.

Knappe, J. (1970). *Annu. Rev. Biochem.* **39**, 757.

Kohler, P. O., Grimley, P. M., and O'Malley, B. W. (1968). *Science* **160**, 86.

Korenman, S. G., and O'Malley, B. W. (1967). *Biochim. Biophys. Acta* **140**, 174.

Landman, A. D., and Dakshinamurti, K. (1973). *Anal. Biochem.* **56**, 191.

Launer, H. F., and Fraenkel-Conrat, H. (1951). *J. Biol. Chem.* **193**, 125.

Lee, H. M., Wright, L. D., and McCormick, D. B. (1973). *Proc. Soc. Exp. Biol. Med.* **142**, 439.

Lichstein, H. C., and Birnbaum, J. (1965). *Biochem. Biophys. Res. Commun.* **20**, 41.

Lynen, F., Knappe, J., Lorch, E., Jutting, G., and Ringelmann, E. (1959). *Angew. Chem.* **71**, 481.

McCormick, D. B. (1965). *Anal. Biochem.* **13**, 194.

McCormick, D. B., and Wright, L. D. (1970). *In* "Methods in Enzymology" (D. B. McCormick and L. D. Wright, eds.), Vol. 18A, p. 379. Academic Press, New York.

Nucleic Acids Research

A tomato cDNA encoding a biotin-binding protein

N.E.Hoffman, E.Pichersky[1] and A.R.Cashmore

Plant Science Institute, Department of Biology, University of Pennsylvania, Philadelphia, PA 19104 and
[1]Laboratory of Cell Biology, The Rockefeller University, New York, NY 10021, USA

We have recently screened a tomato leaf cDNA expression library in the phage vector Charon 16 (1) with antibodies made against chloroplast membrane proteins. The screening procedure utilized secondary antibodies conjugated to biotin and alkaline phosphatase (AP) conjugated to avidin (ABC reagent, Vector Laboratories). The nucleotide sequence of one of the clones obtained and the protein sequence of its longest, but still incomplete, reading frame (Fig. 1) showed no homology to the sequences of interest. A search of protein data banks revealed homology to several biotin-containing proteins from both animals and bacteria, and especially strong homology with the biotinyl subunit of the transcarboxylase of *Propionibacterium* (2). The biotin binding site of the latter, GQTVLVLEAMKME, differs by a single residue from a similar sequence in the tomato protein. Direct screening of phage plaques with avidin-AP in the absence of both primary and secondary antibodies subsequently indicated that this tomato polypeptide directly binds avidin-AP. Our results identify one source of "false" positives in screening cDNA expression libraries with an avidin-biotin based detection system. The tomato cDNA clone, which appears to be the only one identified to date encoding a plant biotin-binding protein, is available upon request.

```
  G   T   V   V   A   P   M   V   G   L   E   V   K   V   L   V   K   D   G   E   K   V   Q
GGTACTGTGGTTGCACCTATGGTTGGGTTAGAGGTTAAAGTATTGGTGAAGGATGGGGAGAAAGTTCAGG    70

  E   G   Q   P   V   L   V   L   E   A   M   K   M   E   H   V   V   K   A   P   A   N   G   Y
AGGGACAACCTGTGTTAGTATTAGAAGCAATGAAGATGGAGCATGTAGTGAAAGCACCAGCTAATGGCTA    140

  V   S   G   L   E   I   K   V   G   Q   S   V   Q   D   G   I   K   L   F   A   L   K   D
TGTAAGCGGGCTTGAAATCAAAGTGGGCCAATCGGTCCAAGATGGTATAAAACTCTTTGCTCTCAAGGAC    210

TGAAATATATCCTGAGGCTATGACAACATCATTCTAGACAACTGTATGTATGACCTCCTATGAGGACTGT    280

ACATCCACTAGGCATAAGAATAACAACATTGAGATGTAAAATCTCTTGGAGTTTCATCTTCTACATTTCA    350

AATACATTTACTTGTAAATGACTTTCCAGACTCA_n                                    384
```

Figure 1: Nucleotide sequence of the tomato cDNA clone and sequence of the encoded polypeptide. The putative biotin-binding site is overlined.

References:
1. DellaPenna, D., et al., Proc. Natl. Acad Sci. USA. *83*, 6420-6424, 1986.
2. Maloy, W.L., et al., J. Biol. Chem. *254*, 11615-11622, 1979.

Tissue- and stage-specific expression of a soybean (*Glycine max* L.) seed-maturation, biotinylated protein

Yue-ie C. Hsing[1], Chih-hua Tsou[1], Tze-fu Hsu[1], Zuey-ying Chen[1], Kim-leong Hsieh[1], Jaw-shu Hsieh[2] and Teh-yuan Chow[1]

[1]*Institute of Botany, Academia Sinica, Taipei, Taiwan (*author for correspondence);* [2]*Department of Agronomy, National Taiwan University, Taipei, Taiwan*

Abstract

A cDNA clone GmPM4 which encodes mRNA species in mature or dry soybean seeds was characterized. DNA sequence analysis shows that the deduced polypeptides have a molecular mass of 68 kDa. GmPM4 proteins have a relatively high amino acid sequence homology with a major biotinylated protein isolated from pea seeds, SBP65, but both of these proteins differ markedly from that of presently known biotin enzymes. The accumulation of *GmPM4* mRNA is detectable in the leaf primodium and the vascular tissues of the hypocotyl-radicle axis of mature seeds, and the GmPM4 proteins are present at high levels in dry and mature soybean seeds, but not in fresh immature seeds. It degrades rapidly at the early stage of seed germination. These proteins are boiling-soluble and biotinylated when they are present endogenously in soybean seeds; however, the same recombinant protein expressed in *Escherichia coli* is boiling-soluble, but it is not biotinylated.

Introduction

Late maturation of seeds is marked not only by water loss but also by a drastic change in profile of proteins synthesized [e.g. 31]. Proteins synthesized during this late stage, which are correlated with desiccation tolerance [1], ABA content [25], or transition to seedling growth [31], are often termed maturation proteins [31] or late embryogenesis-abundant (LEA) proteins [18]. Maturation proteins are slightly different from LEA proteins in that the messages for maturation proteins are not necessarily present at a relatively high level as LEA message during late embryogenesis. Based on the commonly shared amino acid sequence domains, LEA proteins are grouped into three or four groups [8, 9]. Virtually all of the LEA proteins are highly hydrophilic, contain no Cys or Trp residues, and are boiling-soluble [8, 9]. It has been hypothesized that

LEA proteins may play a protective role in plant cell under various stress conditions and this protective role may be essential for the survival of the plant under extreme stress conditions [9, 37].

We have isolated a number of cDNA clones of soybean seed maturation proteins from a pod-dried seed cDNA library by differential screening [23, 24]. These are designated GmPM clones, denoting for *Glycine max* physiologically mature. pGmPM1 [5] and pGmPM9 [22] were found as a member of the group 4 LEA family, and pGmPM2 [19], pGmPM8 and pGmPM10 [20] group 3 LEA family.

In the present study, we report the identification and characterization of a novel soybean seed maturation protein and its cDNA clone, GmPM4. We show that the GmPM4 protein is biotinylated. We investigated the cellular and tissue localization of the RNA transcripts and the expression of this protein during seed development and germination. We also characterized the biochemical behaviors of the recombinant plant GmPM4 proteins in *E. coli*.

Materials and methods

Plant materials

Soybean (*Glycine max* L. cv. Shi-shi) seeds were kindly provided by the Kaohsiung Agricultural Experimental Station, Pintong, Taiwan. Plants were grown to maturity in a field environment. Pods were harvested at mid-development, about 35 days after flowering (DAF), and seeds were precociously matured by air-drying the intact pods (pod-dried, PD) for several days as indicated in each experiment [23, 24]. The fresh weight, dry weight, moisture content and germination behavior of the soybean seeds throughout the course of development, from 21 DAF to 70 DAF, were studied by Rosenberg and Rinne [31]. Germination occurred between 12 h (0% of seed germination) and 24 h (ca. 90% of seed germination) after sowing. Germinating seeds were harvested certain time after imbibition as indicated in each experiment. Soybean seedlings with four compound leaves were also used in the studies. Leaves of control plants or those treated with 2% PEG 6000 for 5 days were harvested. All plant materials were frozen in liquid nitrogen and stored at −70 °C until use.

Isolation of cDNA clones, cDNA sequencing and molecular analysis

A soybean dried-seed cDNA library was constructed and a series of GmPM cDNA clones were selected by differential screening as previously described [24]. The nucleotide sequence of pGmPM4 was determined with the dideoxy chain termination method [33] using Sequenase (United States Biochemicals), and the data were analyzed using the Genetics Computer Group Sequences Analysis software package Version 9.0 [7].

Protein extraction and analysis

Total soluble proteins of soybean seeds or *E. coli* were extracted by homogenizing in ice-cold buffer A consisting of 63 mM Tris-HCl pH 7.8, 20 mM $MgCl_2$ and 1 mM PMSF (phenylmethylsulonyl fluoride). After homogenization, an equivalent volume of Laemmli protein solubilization buffer [28] was added. The sample was incubated at 100 °C for 5 min before gel loading.

Boiling-soluble proteins of soybean seeds or *E. coli* were extracted by homogenizing in ice-cold grinding buffer consisting of 20 mM TES/KOH (pH 8.0) and 500 mM NaCl. The slurry was transferred to a centrifuge tube and incubated at 100 °C for 10 min and then at 4 °C for 10 min. After centrifugation, the proteins in supernatant were concentrated by acetone precipitation. The precipitate was resuspended in Laemmli protein solubilization buffer, and the slurry was then incubated at 100 °C for 5 min before gel loading.

Characterization of GmPM4 cDNA clone was performed by hybrid select translation. Fifty μg of plasmid DNA was denatured and spotted onto nitrocellulose filters. The filters were hybridized with poly(A) RNA prepared from the cotyledons of 4-day PD, 35 DAF soybean seeds. The RNA sample hybridized to the filters was translated *in vitro* using a rabbit reticulocyte lysate system (Promega Biotec, USA) in the presence of ^{35}S-Met.

All protein samples were separated by one-dimensional 12.5% SDS-polyacrylamide gel electrophoresis, and detected by Coomassie blue staining, fluorography or western blot. Biotinylated polypeptides were revealed by AP-conjugated streptavidin and the chromogenic substrate nitroblue tetrazolium. For comparison, the sera against 130 kDa [24] and GmPM8 [20] seed maturation proteins and against seed storage protein glycinin were also used in western blot analysis.

Construction of recombinant plasmid and induction of protein expression

A 2.0 kb *Nde*I-*Eco*RI fragment of pGmPM4 was ligated to the *Nde*I-*Eco*RI fragment of the pET-24a T7 expression vector (Novagen, USA). Transformation of *E. coli*, preparation of plasmid DNA, and other routine procedures were performed according to the protocols of manufacturers and of Sambrook *et al.* [32]. Cells were cultured in LB medium with a supplement of 4 μM biotin. Expression of the GmPM4 recombinant protein was induced by the addition of 1 mM IPTG.

mRNA in situ hybridization

Four-day PD soybean seeds were harvested and immediately fixed with 4% formaldehyde solution. Tissues were fixed and paraffin-embedded according to the procedures described by Jackson [26]. The slides were hybridized with digoxigenin-labeled antisense or sense strands of GmPM4 probes transcribed by T3 or T7 RNA polymerase from linearized pBluescript SK− harboring the GmPM4 cDNA. Labeling with digoxigenin was performed according to instructions provided by the manufacturer (Boehringer

Mannheimm Germany). Hybridization of the probes to the slides was carried out at 50°C overnight, and an anti-digoxigenin-AP conjugate (Boehringer) and the color substrates were used for detection.

In vitro *biotinylation of the recombinant proteins*

About 55 DAF pea (*Pisum sativum*) seeds were used in an in vitro biotinylation reaction of the GmPM4 recombinant protein. These seeds had a mean fresh weight of 450 mg and a water content of 55%. They were ground in buffer A (Tris buffer) or buffer B containing 50 mM HEPES pH 7.4 , 1 mM PMSF, 1 mM EDTA, 1 mM DTT and 10% glycerol at 4°C. The slurry was centrifuged at 11000 $\times g$ and the supernatant was collected as the crude enzyme extract. The in vitro biotinylation assay was carried out in 100 μl reaction mixture containing 50 mM HEPES pH 7.4, 2 mM $MgCl_2$, 1 mM NADH, 2 mM ATP, 2 μM biotin, total *E. coli* proteins containing about 5 μg recombinant GmPM4 protein, and 500 μg of crude enzyme extract from pea seeds. The reaction was carried out at 30°C for 2 h and stopped by adding gel loading buffer. The biotinylation of the proteins was then detected by SDS-PAGE and western blot.

Results

cDNA *cloning and deduced protein sequence analysis*

A hybrid-select translation assay indicated that there is only one protein band corresponding to the GmPM4 cDNA, with an apparent molecular mass of 70 kDa (lane a′, Figure 1a). The nucleotide sequence of pGmPM4 contains 2134 bp with 60 bp and 142 bp at the 5′- and 3′- untranslated region, respectively. The putative protein comprises a 643 amino acids with a predicted molecular mass of 67988 Da, similar to the result obtained from hybrid select translation. The pI value of the deduced GmPM4 protein is 6.1. The protein is highly hydrophilic as revealed by the hydropathy plot (data not shown) and its preponderant amounts of charged amino acids although Ala and Gly are also abundant. Southern blot analysis of the soybean genomic DNA using pGmPM4 insert as the probe reveals one to two intensive bands when hybridized under conditions of high stringency (data not shown). It is therefore possible that the GmPM4 protein is encoded by a small gene family.

Since most of the GmPM clones belong to the LEA family, biochemical properties of the GmPM4
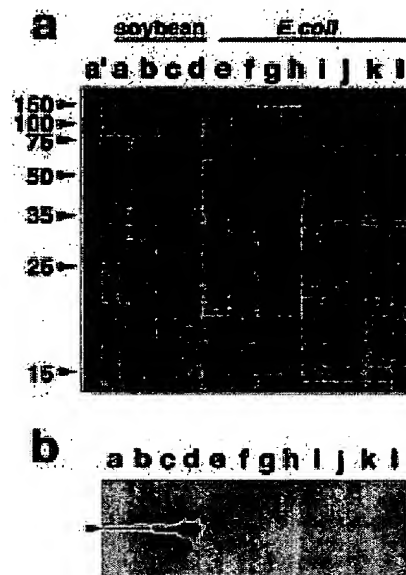
Figure 1. Total proteins and immunoblot analysis of soybean seeds and *E. coli* cells expressing GmPM4. a. Lane a', fluorogram of hybrid-selected translation product using pGmPM4 DNA. Lanes a to 1 are Coomassie blue-stained protein profiles. The following samples were loaded onto the lanes: total proteins (lane a) and boiling-soluble proteins (lane b) from fresh 35 DAF soybean seeds; total proteins (lane c) and boiling-soluble proteins (lane d) from 4-day PD 35 DAF soybean seeds; total proteins (lanes e–h) and boiling-soluble proteins (lanes i–l,) from the control *E. coli* (lanes e, i), the recombinant *E. coli* at 0 h (lanes f, j), 1.5 h (lanes g, k), and 3 h (lanes h, l) after the addition of IPTG to the media. The arrows on the left indicate the molecular weight standards. b. Biotinyl proteins in the crude extract from soybean seeds or *E. coli* using AP-conjugated streptavidin as a specific reagent. Lanes a to l, the same as in Panel a. The arrow on the left indicates the position of 70 kDa.

fragment	length/ res. apart	position	sequences
I	8/26	62-69	ERVKDHAG
			::| |||:
		96-103	DKVTDHAA
II	11/26	268-278	LAAQAKDATLE
			:| ||| ||:
		305-315	VAEKAKDYTLQ
III	10/16	316-325	AAEKAKSAGG
			|:| :|||:|
		342-351	AVESGKSAAG
IV	12/32	349-360	AAGYAAKVAADL
			|||||:: ||:|
		393-404	AAGYAGHKAAEL
V	17/17	556-572	GGEVLGAVGETVAEIGQ
			:| || |:|||:||| :
		590-606	AGGVLDAIGETIAEIAE

Figure 2. Repeated amino acid sequences in GmPM4 protein. The positions, repeat length, and number of residues separating the repeat are indicated.

are compared with those of the LEA proteins. Like many LEA proteins, GmPM4 contain large amounts of Ala, Arg, Glu, Gly, Lys and Thr. However, in contrast to LEA protein, GmPM4 include one Cys and two Trp residues. There are several repeated homologous stretches of 8, 10, 11, 12 or 17 amino acids, as shown in Figure 2. It can be observed that these repeat stretches are separated by several amino acid residues. This character does not share the same property of the group 3 LEA proteins in which direct, tandem repeats of 11-mer are found. The homology scan by BLAST or TFASTA of the Genetics Computer Group Sequence Analysis software package [7] shows a 20 to 25% identity in the central region with about 250 amino acid long of GmPM4 protein with several of the group 3 LEA proteins, including birch BP8 [29], soybean GmPM2 [19], GmPM8 and GmPM10 [20].

GmPM4 protein is similar to pea SBP65

A search for protein sequence homology in data banks reveals a strong similarity between the GmPM4 protein and the biotinylated protein SBP65 from pea seeds [11]. The similarity between the two proteins is 71.5% and the identity 54.1%. The SBP65 is a little smaller, with only 551 amino acids. This protein has been studied intensively by Duval et al. [10–14]. It has an apparent molecular mass of 65 kDa, does not exhibit any of the biotin-containing carboxylase activities and it represents a major part of the total protein-bound biotin in pea seeds [12]. The protein is localized in cytosol [14] and specifically expressed in the late maturating seeds, and is absent in leaf, root, stem, pod and flower tissues of the pea plants [10, 13]. It is proposed that SBP65 acts as a storage form of biotin to support seedling growth during germination, in which a strong demand for biotin-containing carboxylases such as acetyl CoA carboxylase is obviously needed [12].

Figure 1 shows the profiles of total extractable-proteins (Panel a) and avidin detectable biotinylated proteins (Panel b) in several soybean tissues. For detection of biotinylation, a protein blotting and strep-tavidin/alkaline phosphatase (AP) system was used. Only one biotinylated protein was detected in the 4-day PD, 35 DAF soybean seeds, and the apparent molecular mass of this protein was found exactly the same as the GmPM4 protein, i.e. 70 kDa. Some soybean biotin-containing carboxylases had been identified, purified or the cDNA cloned, and the molecular mass of acetyl CoA carboxylase was found to be 58,
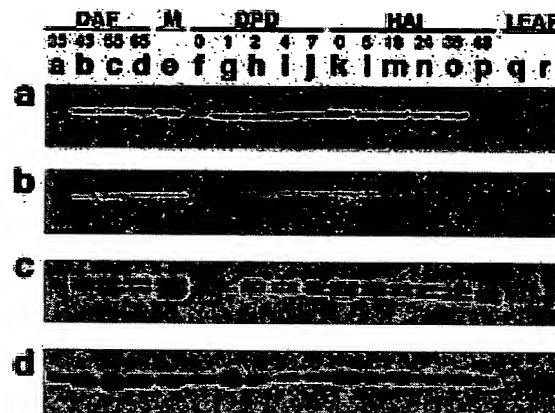
Figure 3. The presence of GmPM4 protein in developing, mature, germinating soybean seeds and seedlings. Biotinyl proteins in the crude extract from soybean seeds were detected using AP-conjugated streptavidin as the specific reagent (Panel a). For comparison, antibodies against 130 kDa soybean seed maturation protein [24] (Panel b), maturation protein GmPM8 [20] (Panel c), and storage protein glycinin (Panel d) were also used. The following samples were loaded onto the lanes: total proteins of 35 DAF (lane a), 45 DAF (lane b), 55 DAF (lane c), 65 DAF (lane d) and mature (M, lane e) soybean seeds, total proteins of 0 day PD (DPD) (lane f), 1 day PD (lane g), 2 day PD (lane h), 4 day PD (lane i), 7 day PD (lane j) 35 DAF soybean seeds, total proteins of 0 h after imbibition (HAI) (lane k), 6 HAI (lane l), 18 HAI (lane m), 24 HAI (lane n), 36 HAI (lane o), 48 HAI (lane p) soybean seeds, and total leaf proteins from control soybean seedlings with four compound leaves (lane q), or those treated with 2% PEG for 5 days (lane r).

65, and 240 kDa [4], and that of methylcrotonoyl-CoA carboxylase was 85 kDa [38]. Hence, the 70 kDa biotinylated protein is very likely the same as GmPM4 protein. This protein is present in very low abundance in the fresh 35 DAF seeds (lane a) but high in the 4-day PD seeds (lane c). This seed protein stays in the supernatant after extracts of the total seed proteins are boiled (lane b and d, Panel a and b), a property similar to most of the soybean seed maturation proteins [1]. We hence have characteristically assigned this soybean GmPM4 protein, together with the pea SBP65 protein, as biotinylated.

Expression of GmPM4 protein in developing and germinating soybean seeds

To evaluate the expression of GmPM4 protein in developing and germinating soybean seeds, we extracted proteins from artificially dried, naturally mature or germinating seeds. Total proteins from young leaves of control soybean seedlings or water-stressed, PEG-treated seedlings were also used in the study. Figure 3a

shows the patterns of biotinylated protein present in these tissues. The data show that little or no GmPM4 protein was detected in young developing seeds (35 DAF, lane a), and the protein accumulates to a high level after 45 DAF (lane b), and remains at a similar high level in 65 DAF (lane d) or mature (lane e) seeds. When the 35 DAF seeds were artificially dried, the GmPM4 protein increased rapidly to a high level after only one day PD treatment (lane g), and remained at high levels in 2 PD (lane h), 4 PD (lane i) and 7 PD (lane j) seeds. The levels of another two soybean seed maturation proteins, 130 kDa [24] (Panel b) and GmPM8 [20] (Panel c), and a storage protein glycinin (Panel d) in these tissues were also studied. These two maturation proteins had the same accumulation patterns as GmPM4 had, i.e. little or no protein was detected in young developing seeds and the levels increased in late maturing or artificially dried seeds. The same pattern of protein accumulation has been reported [1, 31], and is the characteristic of seed maturation proteins. For the storage protein glycinin, however, the amount was already very high in 35 DAF seeds, the youngest one used in the study. After seed imbibition, the levels of GmPM4 and 130 kDa protein decreased gradually and disappeared 48 h after imbibition (HAI, lanes k to p), while the levels of GmPM8 protein and glycinin remain relatively high at 48 HAI. These four seed proteins all appeared to be seed-specific since no protein could be detected in leaves (lane q) and other tissues (data not shown). Also there was no accumulation of these proteins in the leaves of water-stressed seedlings (lanes q and r).

GmPM4 mRNA accumulates in the vascular system

In situ mRNA hybridization analysis with digoxigenin-labeled riboprobes were conducted using transverse sections prepared from 4-days PD 35 DAF soybean seeds. We show that high levels of GmPM4 mRNA are accumulated at leaf primodium (Figure 4c) and at the central stele of hypocotyl-radicle axis (Figure 4b). At the vascular bundle, the level of mRNA is high in the metaxylem, phloem and slightly lower in pith tissues, while no signal detected in the protoxylem tissue (Figure 4d). GmPM4 mRNA was also not detected in the remaining endosperm and the parenchymal cells of cotyledon, and very low messages, if any, in the cortex of hypocotyl-radicle axis and vascular tissues of cotyledon (Figure 4b). Western blot analysis of cotyledon and embryonic axis proteins also revealed that the level of GmPM4 protein was very high in the

axis and low in cotyledon (data not shown). Specificity of in situ hybridization reactions was confirmed by the lack of appreciable reaction of the GmPM4 sense strand probe with the paraffin-embedded sections (Figure 4a).

Based on these results, we conclude that cell-type-specific expression of GmPM4 gene occurs during the late maturation stage of soybean seed development. Gene expression at the transcription level is highest in the leaf primordia and the vascular system of hypocotyl-radicle axis, low in the vascular system of cotyledon, and minimal or not detectable in most other tissues.

Expression of recombinant GmPM4 protein in E. coli

There exists an NdeI site in the multiple cloning site of the expression vector pET-24. The restriction site of NdeI is CATATG, and this ATG should be utilized as the start codon for the recombinant protein. There is only one NdeI site in pGmPM4 DNA, which is located exactly at the start codon. For the expression of GmPM4 protein in E. coli, the 2 kb NdeI-EcoRI fragment from pGmPM4 containing the whole open reading frame (ORF) was ligated to the 5.3 kb NdeI-EcoRI fragment of pET-24a(+). Thus, the recombinant GmPM4 protein utilized its own start and stop codons for translation, and should have the same molecular mass as that in the soybean seeds. Since there was no T7.Tag sequence or His.Tag sequence in the ORF of the recombinant protein, this protein could not be purified using affinity chromatography, as could many other recombinant proteins using the pET systems.

Figure 1 illustrates the effective induction of recombinant GmPM4 protein by the application of IPTG to the culture medium (lanes e to h, Panel a). The apparent molecular mass of the recombinant protein is indeed identical to that from soybean seeds, i.e. 70 kDa as analyzed by SDS-PAGE. This recombinant protein stays in the extraction solution after being boiled (lanes i to l, Panel a) and is thus boiling-soluble as was found for the endogenous protein extracted from soybean seeds. However, the recombinant protein produced from E. coli were not detectable by streptavidin binding assay (lanes f to h, Panel b). Therefore, in contract to the biotinylated soybean GmPM4 produced endogenously, the recombinant GmPM4 protein produced in a heterologous system was not biotinylated in E. coli, although a concentra-
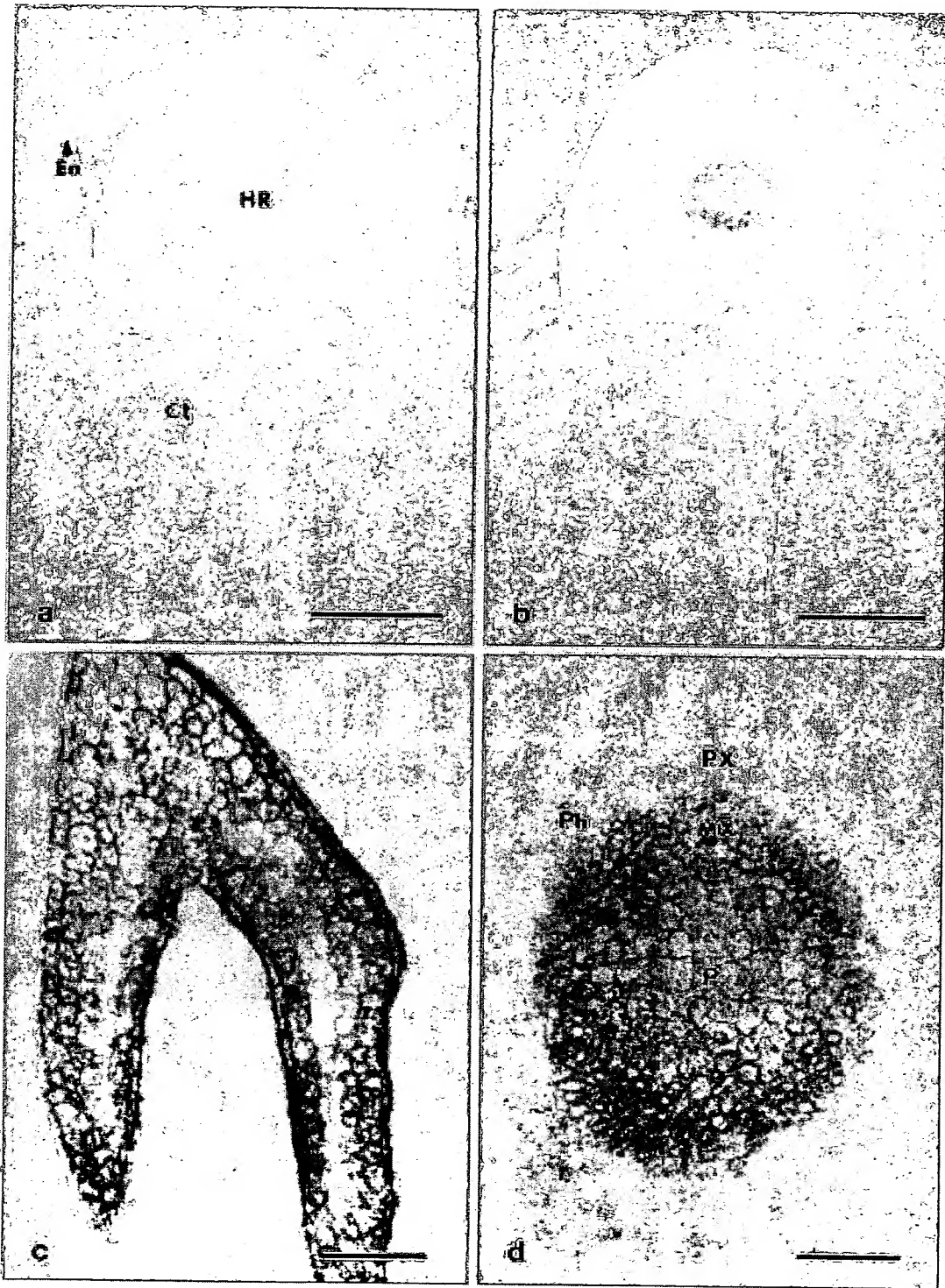
Figure 4. Localization of *GmPM4* transcripts in 4-day PD soybean seeds. *In situ* hybridization was carried out with transverse sections of pod-dried soybean seeds. The development time of the reaction with the color substrates was 25 h. a. Transverse section through the region of embryonic axis and two pieces of cotyledon hybridized with the sense probe; bar = 50 μm. b. Section similar to the one shown in a but hybridized with the antisense probe; bar = 50 μm. c. Cross section of the plumule hybridized with antisense probe; bar = 5 μm. d. Higher magnification of the central stele in b; bar = 10 μm. Abbreviations: Ct, cotyledon; En, endosperm; HR, hypocotyl-radical axis; Mx, metaxylem; P, pith; Ph, phloem; Px, protoxylem.

tion of 4 μM biotin was supplemented to the culture medium.

To test whether the *E. coli* GmPM4 recombinant proteins can be biotinylated by crude seed extracts, we have employed an *in vitro* assay system, where a crude extract was prepared from late developing pea seeds. Since pea SBP65 and soybean GmPM4 proteins have different apparent molecular mass, i.e. 65 and 70 kDa, respectively, we contemplated that this difference may allow us to interpret the test results. NADH, ATP, and biotin were used as energy sources and the substrates in reaction buffer were described in Materials and methods. Figure 5 shows that no 70 kDa biotinylated protein was detected after incubating the recombinant proteins with pea crude extracts for two hours (lanes f and h, Figure 5b). Thus, the GmPM4 recombinant protein obviously could not be biotinylated in the *in vitro* assay employed.

Discussion

A new seed maturation protein group

The developmental pattern of biotinylated proteins (BP) during embryogenesis, maturation and germination of soybean seeds was characterized by Shatters and his colleagues [1, 6, 35]. There are three of these proteins, with an apparent molecular mass of 85 kDa (BP85), 75 kDa (BP75), and 35 kDa (BP35), respectively. These authors predicted that the BP75 might be related to the pea SBP65. Due to the similarities in apparent molecular weight, tissue specificity and protein accumulation patterns, the GmPM4 and BP75 are the same protein.

GmPM4, BP75 and SBP65 proteins were all seed-specific biotinylated proteins present in the embryonic axes and the cotyledons. Only traces of these proteins, if any, were detected at early stages of seed development. The accumulation of these proteins were induced by artificial drying of the immature seeds or by natural maturation which water loss also occurred. After seed imbibition, these proteins disappeared at early stage of germination. These proteins were not present in any other tissues examined including leaves,

Figure 5. *In vitro* biotinylation assay of GmPM4 recombinant proteins. a. Lanes a to i are Coomassie blue-stained protein profiles. The following samples were loaded onto the lanes: total proteins of fresh (lane a), 4-day PD (lane b) 35 DAF soybean seeds, total proteins of late maturing pea seeds extracted with buffer B (lane c) or buffer A (lane d), crude *E. coli* extract expressing GmPM4 (lane i), biotinylation assay including crude *E. coli* extract, crude pea extract (buffer B), and without (lane e) or with (lane f) ATP and NADH, biotinylation assay including crude *E. coli* extract, crude pea extract (buffer A), and without (lane g) or with (lane h) ATP and NADH. Arrows on the left indicate the molecular weight standards. b. Biotinyl proteins detected with AP-conjugated streptavidin. Lanes a to i, the same as in Panel a. Arrows on the left indicate the positions of 65 kDa (pea SBP65) and 70 kDa (soybean GmPM4).

stems or roots. The protein expression patterns were the same as those of seed maturation proteins, and were quite different from other biotin-containing carboxylase.

All the GmPM cDNA clones were pulled out by differential screening. Most of the GmPM clones se-

quenced fulfil the properties of LEA proteins, as reviewed in the Introduction. These GmPM clones represent relatively high abundance of the messages expressed during late seed mature stage. Out of these 456 positive plaques selected as highly expressed messages in pod-dried soybean seeds, the *GmPM6/7* and *GmPM1/9* families accounted for 100 and 46 clones, respectively [24]. However, only 4 of them were GmPM4 clones. The *GmPM4* message is therefore much less abundant than many *LEA* messages.

Several characteristics of the GmPM4 and SBP65 are quite similar to those of LEA proteins, including their being highly hydrophilic, boiling-soluble, and having many A, D, G, K, R and T residues. However, GmPM4 contain 1 Cys and 2 Trp residues and SBP65 contain 2 Cys and 2 Trp residues. This is quite different from the general features of LEA proteins. The amino acid sequences of GmPM4 or SBP65 do not show similarity with group 1, 2 or 4 LEA proteins. Although there are several repeated homologous stretches in both proteins, they are not tandemly-repeating 11-mers as LEA 3 proteins are. The message of *GmPM4* is less abundant, as indicated previously. According to the points mentioned above, soybean GmPM4 and pea SBP65 do not fulfil totally the characteristics of LEA proteins. Nevertheless, they are a new group of seed maturation proteins, and act as biotin storage pool to support early growth during seed germination.

Biotinylated proteins

The coenzyme biotin (vitamin H) is a carrier of carbon dioxide in enzymatic carboxylation and transcarboxylation reactions [e.g. 27]. It is known that there exists a single biotinylated enzyme in *E. coli*, three in yeast, and four in animal or plant systems [3, 15, 30, 39]. The four biotinylated enzymes in plants have been identified as acetyl-CoA carboxylase, 3-methylcrotonoyl-CoA carboxylase, propionyl-CoA carboxylase and pyruvate carboxylase. Several of these carboxylase genes from *E. coli*, yeast, animal and plant had been cloned and sequenced. By the analysis of biotinylated domain and the alignments of the predicted sequences, it was suggested that the biotin is covalently linked to the epsilon amino group of a specific lysine residue located within a highly conserved region of Met-Lys-Met.

Pea SBP65 contains no carboxylase activity, and the analysis of the biotinyl peptide indicated that the biotinyl lysine was located at position 103 [11]. Fig-

```
consensus seq.     V IV      Q  V
of biotin enz.     LXVLEAMKMRXXIXAX
                    |   :   |   :
SBP65 (96-111)     GGVRDMGKFQMESKGG
                   |||||:|||:|  : ||
GmPM4 (118-133)    GGVRDVGKFEMRTEGG
```

Figure 6. Comparisons of the biotinylation domains of biotin enzymes, pea SBP65 and soybean GmPM4. The consensus sequence of biotin enzymes was obtained by alignment of several biotin carboxylases from *E. coli*, yeast, animal and plant. The biotinylated lysine (K) is printed in bold. The double Glys at both ends are underlined.

ure 6 illustrates the biotinylation domains of biotin enzymes, SBP65 and GmPM4. These regions are highly conserved for SBP65 and GmPM4, with double Glys at both ends, three acidic and three basic amino acids in between, and the biotinyl Lys at the center of the conserved motif. The biotinylation of protein is an ATP-dependent process in which biotinyl-AMP is synthesized as an intermediate [6]. The double Gly residues might play a very important role for the structure and function of GmPM4 and SBP65 proteins. Previous studies have shown that several glycinyl residues are close to a reactive lysinyl residue in a motif involved in nucleotide-binding of several proteins. For instances, a segment of Gly-Gly-Pro-Gly-Ser-Gly-Lys containing the reactive Lys is important for binding of the substrates in adenylate kinase [40], and a segment of Lys-Thr-Gly-Gly-Leu is the active site for starch synthase [17].

Figure 6 shows that only the biotinylated Lys and one Val are conserved comparing the GmPM4 and SBP65 with the consensus sequences of biotin-containing enzymes. Several biotinylated carboxylases from animal or plant, e.g. soybean 3-methylcrotonoyl-CoA carboxylase, were expressed in *E. coli*, and the resulting chimeric protein was biotinylated [38]. In our study, however, the recombinant, *E. coli*-produced GmPM4 protein was not biotinylated (Figure 1). Thus, although the biotin-carrying domain of biotin-containing carboxylase is conserved throughout all kingdoms, the domain of SBP65 and GmPM4 proteins appears to be unique and could not be recognized by a prokaryotic biotinylation system.

Possible roles of GmPM4 proteins in soybean seed development

It has been suggested that the presence of soybean seed maturation proteins might be a prerequisite for normal germination and the subsequent seedling growth [21, 31]. Using pea as a model system, Duval *et al.* [9]

demonstrated that there was a very high demand for biotin in early seedling growth since the activities of acetyl CoA carboxylase and 3-methylcrotonoyl-CoA carboxylase increased strongly in the seedlings 3 to 8 days after imbibition. Since the GmPM4 proteins do not appear to fulfil the biochemical properties of LEA proteins, their messages are not very abundant in mature seeds and will not express in water-stressed seedlings, we suggest that the physiological roles of GmPM4 protein might differ from those of the LEA proteins, i.e. desiccation protection. Instead, these proteins are required for the subsequent normal germination since they would supply an essential cofactor, biotin, for seedling growth.

The *GmPM4* message is tissue-specific and localized to leaf primodium or the vascular tissues of the cotyledon and the hypocotyl-radicle axis. The reserve hydrolysis in germinating seeds has specific patterns; for instance, the mobilization begins around vascular strands for germinating soybean seeds [2]. The localization of GmPM4 proteins and the initiation site of reserve utilization coincide well and indicate that the possible physiological roles for the biotin ligands in GmPM4 proteins might be utilized for relating carboxylases during early stage of seed germination. The importance of biotin in seed development was also demonstrated by the identification and characterization of a biotin auxotroph mutant of *Arabidopsis* [34]. No biotin was detectable in the seeds of this embryo-lethal mutant, but these embryos could be rescued when grown in the presence of biotin [36]. The presence of GmPM4 protein in mature soybean seeds might be essential to function as a source of biotin for several growth-limiting enzymes that are necessary during seed development and the subsequent germination stages, and thus may play some roles in determining seed germination capacity.

Acknowledgements

References

1. Blackman SA, Wettlaufer SH, Obendorf RL, Leopold AC: Maturation proteins associated with desiccation tolerance in soybean. Plant Physiol 96: 868–874 (1991).
2. Bewley JD, Black M: Physiology and Biochemistry of Seeds, vol. 1. Development, Germination and Growth, p. 233. Springer-Verlag, Berlin/Heidelberg/New York (1983).
3. Chandler CS, Ballard FJ: Regulation of the breakdown rates of biotin-containing proteins in Swiss 3T3-L1 cells. Biochem J 251: 749–755 (1988).
4. Charles DJ, Cherry JH: Purification and characterization of acetyl-CoA carboxylase from developing soybean seeds. Phytochemistry 25: 1067–1071 (1986).
5. Chen Z, Hsing YC, Lee P, Chow T: Nucleotide sequences of a soybean cDNA encoding an 18 kilodalton late embryogenesis abundant protein. Plant Physiol 99: 773–774 (1992).
6. Cronan JE Jr: The *E. coli* operon: transcriptional repression by an essential protein modification enzyme. Cell 58: 427–429 (1989).
7. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. Nucl Acids Res 12: 387–395 (1984).
8. Dure L III: Structural motifs in Lea proteins. In: Close TJ, Bray EA (eds) Plant Responses to Cellular Dehydration during Environmental Stress. Current Topics in Plant Physiology, vol. 10, pp. 91–103. American Society of Plant Physiologists Series (1993).
9. Dure L III, Crouch M, Harada J, Ho T-hD, Mundy J, Quatrano RS, Thomas T, Sung ZR: Common amino acid sequence domains among the LEA proteins of higher plants. Plant Mol Biol 12: 475–486 (1989).
10. Duval M, DeRose RT, Job D, Faucher D, Douce R, Job D: The major biotinyl protein from *Pisum sativum* seeds covalently binds biotin at a novel site. Plant Mol Biol 26: 265–273 (1994).
11. Duval M, Job C, Alban C, Douce R, Job D: Developmental patterns of free and protein-bound biotin during maturation and germination of seeds of *Pisum sativum*: characterization of a novel seed-specific biotinylated protein. Biochem J 299: 141–150 (1994).
12. Duval M, Job C, Alban C, Sparace S, Douce R, Job D: Synthesis and degradation of a novel biotinyl protein in developing and germinating pea seeds. C R Acad Sci (Paris), Ser III 316: 1463–1470 (1993).
13. Duval M, Loiseau J, Deuaye L, Pepin R, Ledeunff Y, Wang T, Job D: SBP65, a seed-specific biotinylated protein, behaves as a LEA protein in developing pea embryos. C R Acad Sci (Paris) Ser III 319: 585–594 (1996).
14. Duval M, Pepin R, Job C, Derpierre C, Douce R, Job D: Ultrastructural localization of the major biotinylated protein from *Pisum sativum* seeds. J Exp Bot 46: 1783–1786 (1995).
15. Fall RR: Analysis of microbial biotin proteins. Meth Enzymol 67: 390–398 (1979).
16. Franca Neto JB, Shatters RG, West SH: Developmental pattern of biotinylated proteins during embryogenesis and maturation of soybean seed. Seed Sci Res 7: 377–384 (1997).
17. Furukawa K, Tagaya M, Masayori I, Preiss J, Fukui T: Identification of Lysine 15 at the active site in *Escherichia coli* glycogen synthase. J Biol Chem 265: 2086–2090 (1990).
18. Galau GA, Hughes DW, Dure L: Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. Plant Mol Biol 7: 155–170 (1986).
19. Hsing YC, Chen Z, Chow T: Nucleotide sequences of a soybean complementary DNA encoding a 50-kilodalton late

490

embryogenesis abundant protein. Plant Physiol 99: 354–355 (1992).

20. Hsing YC, Chen Z, Shih M, Hsieh J, Chow T: Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds. Plant Mol Biol 29: 863–868 (1995).

21. Hsing YC, Hsieh K, Huang Y, Hsieh J: Premature drying and germination in wild soybean seeds. Taiwania 40: 73–81 (1995).

22. Hsing YC, Lee P, Chen Z, Chow T: A soybean seed cDNA (X63565) encoding a late embryogenesis abundant protein. Plant Physiol 109: 1125 (1995).

23. Hsing YC, Rinne RW, Hepburn AG, Zielinski RE: Expression of maturation-specific genes in soybean seeds. Crop Sci 30: 1343–1350 (1990).

24. Hsing YC, Wu S: Cloning and characterization of cDNA clones encoding soybean seed maturation polypeptides. Bot Bull Acad Sin 33: 191–199 (1992).

25. Hughes DW, Galau GA: Developmental and environmental induction of Lea and LeaA mRNAs and the post-abscission program during embryo culture. Plant Cell 3: 605–618 (1991).

26. Jackson DP: In situ hybridization in plants. In: Bowles DJ, Gurr SJ, McPhereson M (eds) Molecular Plant Pathology: A Practical Approach, pp. 163–174. Oxford University Press, London (1992).

27. Knowles JR: The mechanism of biotin dependent enzymes. Annu Rev Biochem 58: 195–221 (1989).

28. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685 (1970).

29. Puupponen-pimia R, Saloheimo M, Vasara T, Ra R, Gaugecz J, Kurten U, Knowles JKC, Keranen S, Kauppinen V: Characterization of a birch (Betula pendula Roth.) embryogenic gene, BP8. Plant Mol Biol 23: 423–428 (1993).

30. Robinson BH, Oei J, Saunders M, Grvel R: ^{3}H-Biotin-labeled proteins in cultured human skin fibroblasts from patients with pyruvate carboxylase deficiency. J Biol Chem 258: 6660–6664 (1983).

31. Rosenberg LA, Rinne RW: Moisture loss as a prerequisite for seedling growth in soybean seeds (Glycine max (L.) Merr.). J Exp Bot 37: 1663–1674 (1986).

32. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

33. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 64: 5463–5467 (1977).

34. Schneider T, Dinkins R, Robinson K, Shellhammer J, Meinke DW: An embryo-lethal mutant of Arabidopsis thaliana is a biotin auxotroph. Devel Biol 131: 161–167 (1989).

35. Shatters RG, Boo SP, Franca Neto JB, West SH: Identification of biotinylated proteins in soybean (Glycine max L.) seeds and their characterization during germination and seedling growth. Seed Sci Res 7: 373–376 (1997).

36. Shellhammer J, Meinke D: Arrested embryos from the biol auxotroph of Arabidopsis thaliana contain reduced levels of biotin. Plant Physiol 93: 1162–1167 (1990).

37. Skriver K, Mundy J: Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2: 503–512 (1990).

38. Song J, Wurtele ES, Nikolau BJ: Molecular cloning and characterization of the cDNA coding for the biotin-containing subunit of 3-methylcrotonoyl-CoA carboxylase: Identification of the biotin carboxylase and biotin-carrier domains. Proc Natl Acad Sci USA 91: 5779–5783 (1994).

39. Wurtele ES, Nikolau BJ: Plants contain multiple biotin enzymes: discovery of 3-methylcrotonoyl-CoA carboxylase, propionyl-CoA carboxylase and pyruvate carboxylase in the plant kingdom. Arch Biochem Biophys 278: 179–186 (1990).

40. Yoneya T, Tagaya M, Kishi F, Nagazawa A, Fukui T: Site-directed mutagenesis of Gly-15 and Gly-20 in the glycine-rich region of adenylate kinase. J Biochem 105: 158–160 (1989).

Molecular cloning and nucleotide sequence
of chicken avidin-related genes 1–5

Riitta A. KEINÄNEN[1], Mika J. WALLÉN[2], Paula A. KRISTO[2], Mikko O. LAUKKANEN[1], Tarja A. TOIMELA[1], Merja A. HELENIUS[1] and Markku S. KULOMAA[1,3]

[1] University of Tampere, Department of Biomedical Sciences, Finland
[2] Research Laboratory Ltd Alko, Helsinki, Finland
[3] University of Jyväskylä, Department of Biology, Finland

Using avidin cDNA as a hybridisation probe, we detected a gene family whose putative products are related to the chicken egg-white avidin. Two overlapping genomic clones were found to contain five genes (avidin-related genes 1–5, avr1–avr5), which have been cloned, characterized and sequenced. All of the genes have a four-exon structure with an overall identity with the avidin cDNA of 88–92%. The genes appear to have no pseudogenic features and, in fact, two of these genes have been shown to be transcribed. The putative proteins share a sequence identity of 68–78% with avidin. The amino acid residues responsible for the biotin-binding activity of avidin and the bacterial biotin-binding protein, streptavidin, are highly conserved. Since avidin is induced in both a progesterone-specific manner and in connection with inflammation, these genes offer a valuable tool to study complex gene regulation in vivo.

Avidin is a basic glycoprotein found in avian, reptilian and amphibian egg white (Hertz and Sebrell, 1942; Jones and Briggs, 1962; Korpela et al., 1981). It is composed of four identical subunits, each consisting of 128 amino acids, of known sequence (DeLange and Huang, 1971). Avidin is capable of binding the vitamin biotin with an exceptionally tight non-covalent bond ($K_a = 10^{-15}$ M; Green, 1975), suggesting that it may function as an antibiotic protein inhibiting bacterial growth.

Avidin is induced in the oviduct of the estrogen-pretreated chick by a single steroid hormone, progesterone (O'Malley et al., 1969; Tuohimaa et al., 1989). This progesterone induction requires mRNA synthesis and is followed by a parallel increase in both the avidin mRNA and protein amounts suggesting that the induction occurs mainly at the transcriptional level. In addition, avidin is induced in connection with tissue trauma (Heinonen et al., 1978) and toxic agents (Elo et al., 1975; Heinonen and Tuohimaa, 1978) as well as a consequence of heat injury (Elo, 1980) and bacterial or viral infections (Elo et al., 1980; Korpela et al., 1982; Korpela et al., 1983). This expression is detected in all chicken tissues tested except the brain. Since the inflammation-associated induction needs no estrogen pretreatment and can be abolished by anti-inflammatory drugs (Nordback et al., 1982; Niemelä, 1986), the product is called 'inflammation avidin'. Altogether, avidin seems to have at least two partially different induction mechanisms (Tuohimaa et al., 1989).

In order to isolate the avidin gene, the cDNA was cloned (Gope et al., 1987). This cDNA was used as a hybridisation probe in the isolation of a genomic clone (λgAV1, previously called λgAV12201; Keinänen et al., 1988). The λgAV1 clone contains three genes closely related to the avidin cDNA. We call these genes avidin-related genes 1–3, avr1–avr3 (referred to as pgAV1.8, pgAV3.7 and pgAV3.3 in Keinänen et al., 1988), since they are unable to encode the known peptide sequence of egg-white avidin (DeLange and Huang, 1971). In this study, we report molecular cloning of two more members of the chicken avidin gene family and the nucleotide sequence of all five avidin-related genes.

MATERIALS AND METHODS

Materials

Restriction and modifying enzymes were purchased from New England Biolabs, Boehringer-Mannheim and Promega. The Sequenase version 2.0 or Sequenase version 1.0 DNA sequencing kit was a product of United States Biochemicals. The AutoRead fluorescent sequencing kit and the Fluoro-Prime kit were purchased from Pharmacia. Nylon filter (Hybond-N, 0.45-μm pore size) was obtained from Amersham International and [α-³⁵S]dATP, [α-³²P]dTTP and [α-³²P]dCTP (3000 Ci/mmol) were obtained from NEN Research Products. Common laboratory reagents were from Sigma, Merck or J. T. Baker.

Construction and screening of a chicken oviduct genomic DNA library

A chicken genomic library was constructed as previously described (Kleinsek et al., 1986) and was kindly provided by

Prof. B. W. O'Malley. Briefly, chicken oviduct DNA was partially digested by *Mbo*I. 16–24-kb fragments were separated and ligated into the *Bam*HI-digested bacteriophage derivative EMBL4 vector before packaging *in vitro*. Genomic clones were screened using nick translated (Rigby et al., 1977) ³²P-labelled chicken avidin cDNA (Gope et al., 1987) probes, either the full-length cDNA or its 3′ end (397 bp; positions 198–594), in *in situ* plaque hybridisation (Benton and Davis, 1977; Keinänen et al., 1988). The *Escherichia coli* NM539 strain was used to provide host cells.

Characterization of the genomic clones

Recombinant DNA from the genomic clone λgAV2 was isolated for characterization according to Benson and Taylor (1984). The insert was subcloned (Maniatis et al., 1982) into the pBR322 plasmid and an M13 sequencing vector after *Hind*III digestion. Recombinant plasmids were introduced into competent *E. coli* RR1 cells (Hanahan, 1983).

Plasmid DNA was isolated from the transformants by an alkaline-detergent method (Birnboim and Doly, 1979) and digested with *Hind*III. After transfer of DNA onto a nylon filter (Southern, 1975), the inserts were characterized by DNA hybridisation analysis. For hybridisations, 20×NaCl/Cit was prepared containing 3.0 M NaCl, 0.3 M sodium citrate, pH 7.0. Reagents for 50×Denhardt's solution were 5 g Ficoll (type 400; Pharmacia), 5 g poly(vinylpyrrolidone), 5 g bovine serum albumin (fraction 5; Sigma) and H₂O to 500 ml. Prehybridisation was performed in 6×NaCl/Cit/10×Denhardt's solution/0.1% SDS/50 μg/ml denatured herring sperm DNA at 68°C for at least 4 h with hybridisation overnight at 68°C in an identical solution containing the denatured ³²P-labelled cDNA probe [full-length probe, or the 5′-end (positions 1–197), or the 3′-end (positions 198–569) of the cDNA]. Filters were washed at 68°C with 2×NaCl/Cit/0.1% SDS (3×5 min and 3×30 min) and with 0.1×NaCl/Cit/0.1% SDS (30 min) before autoradiography. In addition, subclones giving positive hybridisation signals were analysed by restriction enzyme mapping and sequencing.

For genomic blotting, the chromosomal DNA was isolated from chicken muscle (Strauss, 1988) and digested with *Eco*RI, *Bam*HI or *Hind*III. DNA was analysed on an 0.8% agarose gel. The DNA hybridisation analysis was performed as described above using ³²P-labelled full-length cDNA as a probe.

Polymerase chain reaction amplification and cloning of the avr3 gene

To obtain the 5′ end of the *avr3* gene, primers specific for the *avr* genes and the avidin cDNA (corresponding to nucleotide positions 202–222 and 1302–1321 in Fig. 3) were used to amplify the genes from chicken genomic DNA. As template, 100 ng DNA was used in 50-μl reactions with buffer A (final concentration 50 mM KCl, 10 mM Tris/HCl, pH 9.0, 25°C, with 1.5 mM MgCl₂, 0.01% gelatin, 0.1% Triton X-100), 50 pmol both primers, 200 μmol of each dNTP and 2.5 U *Taq* DNA polymerase overlayed with 100 μl mineral oil. Amplification reactions took place in a thermal cycler (PTC-100 programmable thermal controller, MJ Research). Thirty cycles of amplification (each cycle including denaturation at 94°C for 1 min, primer annealing at 55°C for 1 min and polymerisation at 72°C for 1.5 min) were performed.

The polymerase-chain-reaction (PCR) products were purified from NuSieve GTG agarose (FMC BioProducts) using Magic PCR Preps (Promega) and cloned using a TA Cloning System version 1.3 (Invitrogen). Plasmid DNA was isolated with Magic Minipreps (Promega). The clones for sequencing were selected by restriction enzyme analysis and large-scale isolation was performed with Magic Maxipreps (Promega).

DNA sequencing

The nucleotide sequence was determined by the dideoxy-nucleotide chain-termination sequencing method (Sanger et al., 1977) using modified T7 DNA polymerase. The Klenow fragment of DNA polymerase I was also used for sequencing of *avr1*. Double-stranded sequencing of the PCR-generated fragment of *avr3* was performed following the instructions of the AutoRead fluorescent sequencing kit (Pharmacia). Analysis of the reaction products was made using a Pharmacia ALF automated DNA sequencer (Ansorge et al., 1987). The primers were either M13 universal or *avr*-specific primers synthesised with a Cyclone DNA synthesiser (Milligen/Bioscarch). For automated sequencing, M13 universal primers or *avr*-specific primers were synthesised with the aid of an Applied Biosystems 381A DNA synthesiser and fluorescence-labelled by the FluorePrime kit from Pharmacia. All sequences were determined from both strands.

Data analysis

Sequence data were analysed using programs of the Genetics Computer Group package, version 7.0 (University of Wisconsin Genetics Computer Group, Madison, WI, USA) installed on a Convex C3840 computer (Devereux et al., 1984).

RESULTS

Based on the hybridisation analysis of the chicken genomic DNA, a single copy gene for avidin was suggested (Gope et al., 1987). Since three different genes closely related to the avidin cDNA had already been cloned (Keinänen et al., 1988), the hybridisation analysis of genomic DNA was repeated (Fig. 1). Instead of the two bands observed previously (Gope et al., 1987), three bands were detected in the *Hind*III-digested genomic DNA using the full-length avidin cDNA as a probe. The largest DNA fragment was estimated to be 4.5–5.0 kb. Hybridisation signals obtained from the 3.5–4.0-kb and 2.7–2.9-kb fragments were more intense than signals obtained from the largest fragment, thus suggesting the presence of more than one positive fragment of nearly equal size in both bands. The result confirmed the presence of multiple genes for avidin. In order to isolate the entire avidin gene family, screening of the genomic library was continued.

The second genomic clone, λgAV2, was detected after screening of 3.5×10^4 more clones from the genomic library (Kleinsek et al., 1986). By restriction enzyme and hybridisation analyses the two clones were found to be partially overlapping. The insert of λgAV2, approximately 20 kb, was digested with *Hind*III and subcloned into pBR322. In addition to *avr1*–*avr3* (Keinänen et al., 1988), two more members of the gene family, *avr4* and *avr5*, were isolated.

Since the insert of λgAV1 was subcloned as *Eco*RI fragments (Keinänen et al., 1988), the *avr3* gene containing an
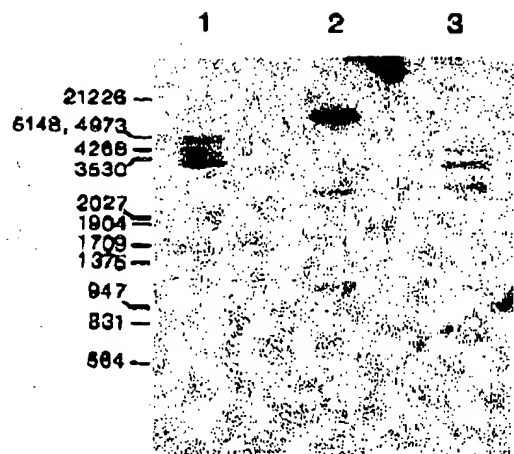
Fig. 1. Hybridisation analysis of chicken genomic DNA. Chromosomal DNA was isolated from chicken muscle, digested with *Eco*RI, *Bam*HI or *Hind*III, electrophoresed on an 0.8% agarose gel and hybridised with the full-length avidin cDNA probe as described in the Materials and Methods section (*Eco*RI, lane 1; *Bam*HI, lane 2; *Hind*III, lane 3). Molecular-size markers (λ/*Eco*RI, *Hind*III) are shown (bp).

internal *Eco*RI site was cut. In order to obtain the 5′ end of the gene, PCR was performed. Primers, corresponding to nucleotide positions 202–222 and 1302–1321 in Fig. 3, were used to amplify a fragment of 1.1 kb from the chromosomal DNA, as expected. Based on restriction enzyme analysis of the produced subclones, one of these clones was assumed to contain the region upstream of the *Eco*RI site of *avr3*.

Different methods were used to determine the nucleotide sequences of the *avr* genes. Both strands were sequenced at least once. The length, approximately 1.1 kb, and general structure of the five genes was similar (Fig. 2). Although sequences of the *avr4* and *avr5* genes are with a single exception identical, on the basis of restriction enzyme mapping of the *Hind*III clones (3.9 kb), we consider them to be different genes. All five genes contained four putative exons split by three introns which follow the GT-AG rule at the predicted exon-intron junctions. Identity between the putative exons and the avidin cDNA was 77–98%. The amino acid sequences deduced from the nucleotide sequences of *avr1*–*avr5* showed an identity of 68–78% when compared with avidin. Identities of single exons with those of preavidin were 57–100% (Table 1).

DISCUSSION

Previous hybridisation studies suggested that avidin was a single-copy gene (Gope et al., 1987). The nucleotide sequence of the five avidin-related genes (Fig. 3) reveals that all of these cloned genes have a single *Hind*III site close to the 3′ end. Thus, the identification in this study of three positive bands from the *Hind*III-digested genomic DNA, two of which are suggested to be at least duplets (Fig. 1), is in keeping with current knowledge of the avidin gene family. Hybridisations of the genomic DNA were performed using the 3′ end (384 bp: Gope et al., 1987) and the full-length cDNA in this study as a probe, but this does not explain the discre-
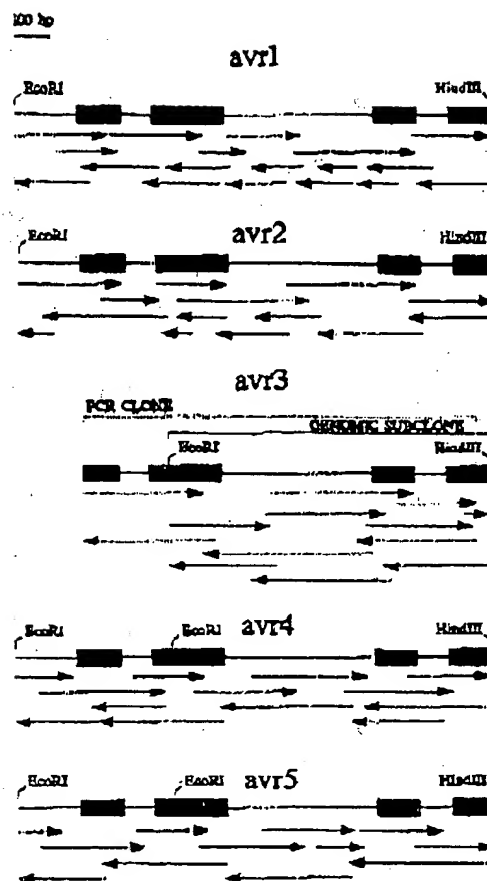
Fig. 2. Sequencing strategy for avidin-related genes 1–5. The exons (■) and sequences obtained from the PCR clone containing the 5′ end of the *avr3* gene (–·) are shown. Sequences from the genomic subclones are also indicated.

Table 1. Identity between the peptide chain of preavidin and putative avidin-related proteins. The identities are shown for each exon and between the mature avidin (without the signal peptide) and putative avidin-related proteins.

Protein	Identity with				
	exon 1	exon 2	exon 3	exon 4	mature protein
	%				
AVR1	100	68	73	64	70
AVR2	96	66	73	57	68
AVR3	96	69	78	64	72
AVR4/5	96	69	95	64	78

pancy between the earlier and now reported results. The genomic DNAs were, however, of different origin. In the earlier experiment, DNA was isolated from chicken spleen and in this study DNA was isolated from chicken muscle. It would,

Fig. 3.

Fig. 3. Sequence alignment of avidin-related genes 1-5 and avidin cDNA. The hyphens represent the corresponding nucleotide for the consensus sequence; nucleotides that differ from the consensus sequence are shown in small letters. The hyphens on the consensus sequence indicate that no consensus was found. Gaps in the sequence are indicated (·). The actual sequence sizes are 1335 nucleotides (avr1 and avr2) due to a gap at nucleotide positions 531-536 in the figure, 1133 nucleotides (avr3) and 1334 nucleotides (avr4 and avr5), the latter three genes having gaps at positions 531-536 and position 1011. The sequence of avr3 starts at position 202 and ends at nucleotide 1321. The first and last nucleotides of the cDNA corresponding to putative exons 1-4 in the avr genes are numbered according to the previously published sequence (Gope et al., 1987) in parenthesis.

therefore, be interesting to further study the methylation pattern or other modifications of the corresponding genes in these tissues to see whether they could explain the different results on the complexity of the avidin-encoding nucleotide sequences in the genomic DNA.

The presence of the avr genes in two partially overlapping genomic clones indicates that they lie in the same DNA cluster. The exon-intron junctions of the avr genes agreed well with the general splice junction consensus sequences (Senapathy et al., 1990). Moreover, sequences which are in accordance with the standard branch point consensus sequence were found at expected intron regions, suggesting that the putative transcripts could be properly processed. The presence of a polyadenylation signal (AATAAA) in all genes suggests that the correct modification of the 3' ends could occur. Finally, no internal stop codons were detected in the putative exon sequences. Altogether, this is consistent with the concept that the avr genes are not pseudogenes and could be transcribed into mature mRNAs. In fact, we have shown that at least two of these genes, avr2 and avr3, are transcribed at a low level in the chicken oviduct and intestine, respectively, after intraperitoneal E. coli infection (Kunnas et al., 1993). In addition, avr3 transcripts have also been detected in a chicken macrophage HD-11 cell line (Lappalainen, P. J., Kunnas, T. A., Punnonen, E.-L. and Kuloma, M. S., unpublished results).

Exon 1, encoding the entire signal peptide of the putative avidin-related proteins (AVR) and the first three amino acids of the mature proteins, was the most fully conserved region of the avr genes were compared with the avidin cDNA. Its identity with the cDNA was 97-98%. Exon 2 of all the

avr genes had regions of clustered point mutations, and the identity of the avr genes with the avidin cDNA was 77-82%. The identity in the third exon was 83-98% and the mutations were again clearly clustered. The fourth exon was also well conserved (92-94%). The point mutations in the introns were more uniformly distributed than in the exons. There was a 6-bp deletion in all avr genes at nucleotide positions 531-536 (exon 2) and a single gap was observed in avr3-avr5 at nucleotide position 1011 (intron 2), but no other deletions or insertions (Fig. 3). The 5' flanking region (up to position -176 from the predicted transcription-initiation site) of avr1, avr2, avr4 and avr5 was perfectly conserved (except for the difference at nucleotide position 10 in Fig. 3), which might predict the expression of the genes, i.e. the promoter region has been subject to evolutionary pressure.

The 6-bp deletion in exon 2 would indicate a common ancestor for the avr genes and might be due to 'slipped-strand' mispairing based on the surrounding repeat-like region containing only cytosine and adenine residues. The distribution of point mutations between the avr genes was markedly similar, indicating a relatively recent duplication of the ancestral gene. Moreover, avr1 and avr2 have many common nucleotides at the sites where they differ from avr4 and avr5, which, as mentioned above, were identical except for the guanine/adenine transition (nucleotide at position 10 in Fig. 3). This would indicate an initial diversion between two genes, one being the ancestor for avr1 and avr2. The ancestor for avr4 and avr5, and possibly for avr3, as indicated by a dendrogram (data not shown) produced by the program PILEUP of the GCG package (Devereux et al., 1984), has

```
                    (1      signal peptide    24)
                                           1                          26
preAVR1    MVHATSPLLL LLLLSLALVA PgLSARKCSL IGKWdNDLGd IMTIGAVNdn
preAVR2    MVHATSPLLL LLLLSLALVA PSLSARKCSL IGaWdNDLGd IMTIGAVNdn
preAVR3    MVHttTSPLLL LLLLSLALVA PSLSARKCSL TGKWINdLGd IMTIrAVNsR
preAVR4/5  MVHttTSPLLL LLLLSLALVA PSLSARKCSL TGKWINdLGd IMTIrAVNsR
preAV      MVHATSPLLL LLLLSLALVA PgLSARKCSL TGKWINDLGd ARTIGAVNsR
SAV        .......... .......... ....Aaeagi TGtWyNqLGd tfivtAg.ad

           27                                              71
preAVR1    GEFnCTNITA VADNPCN.IT rSPLLGIQH. .K..RAcQPT FGFTVHRN..
preAVR2    GEFdGTNITA VADNPCN.IT LSPLLGIQH. .K..RASQPT FGFTVHRN..
preAVR3    GEFaCTNITA VADNPCN.Ik LSPLICIQH. .K..RAcQPT FGFTVHRR..
preAVR4/5  GEFTGTNITA VADNPCN.IT LSPLLGIQH. .K..RASQPT FCFTVHRR..
preAV      GEFTGTNITA VtatEne.Ik eSPLhGtQnt in..krtQPT FGFTVRRK..
SAV        GaLTGTNISA Vgn....aes ryvLtGryds apatdgSgta lGwTVawkna

           72                                             119
preAVR1    FT.BRSTAVT VGQCFVDkSG KEVLKTNRLQ RLAVDCISDD WKATRVGrNd
preAVR2    FT.BRSTAVT VCQCFVDRSG KEVLKTNRLQ RLAVCDISDD WiATRVGnNd
preAVR3    FT.BRSTAVT VGQCFIDRSG KEVLKTNRLQ RLAVDDISDD WKATRVGyNn
preAVR4/5  FT.BRSTYTT tCQCFIDRnG KRVLKTaRLl RSsVnDIsyD WKATRVGyRn
preAV      FD.BRSTYTT tCQCFIDRnG KEVLKTaRLL RSsVnDIgDD WKATRVGlNd
SAV        ynahBATDGS sGGYv..gga earinTgRLL tSgtte.ana WKsTlVGHdt

           128
preAVR1    FTRORTVEE
preAVR2    PTRQhTVEE
preAVR3    FTRORTVEE
preAVR4/5  FTRlaTVEE
preAV      FTRLRTqkE
SAV        FTkvkpsaa
```

Fig. 4. Amino acid sequence of the putative avidin-related preproteins (preAVR1-5), preavidin (preAV) and streptavidin (SAV). Amino acids important for the biotin binding of avidin and streptavidin and corresponding amino acids in the avr genes are shown in boxes. Avidin and streptavidin have been aligned according to Livnah et al. (1993). The numbering refers to the avidin sequence.

supposedly duplicated more recently than the ancestor for avr1 and avr2.

Based on biochemical studies, certain amino acid residues are known to be important for the function of avidin and the bacterial biotin-binding protein streptavidin (Gitlin et al., 1987; Gitlin et al., 1988a; Gitlin et al., 1988b; Kurzban et al., 1989; Weber et al., 1989; Gitlin et al., 1990). Solving the three-dimensional structure of avidin and its functional complex with biotin revealed the amino acid residues responsible for biotin binding (Livnah et al., 1993; Pugliese et al., 1993). Of the 17 residues, 12–14 are conserved in all AVR species (Fig. 4). Thr38, Ala39 and Thr40, forming hydrogen bonds with the valeryl moiety of biotin, are substituted in the AVR species. However, these amino acids are also substituted in streptavidin, which folds very similarly with avidin (Weber et al., 1989; Pugliese et al., 1993). Out of the three lysine residues suggested earlier (Gitlin et al., 1987), Lys111 is most likely part of the biotin-binding site (Pugliese et al., 1993). It is conserved in all AVR species except AVR2. The carbohydrate chain of avidin is attached to Asn17, which in the putative products of avr genes was substituted by isoleucine. The AVR species, however, contained 2–4 potential glycosylation sites (Asn–Xaa–Thr/Ser), although the carbohydrate side chain has been suggested to be insignificant for the biotin binding of avidin (Hiller et al., 1987). Secondary-structure prediction (Chou and Fasman, 1978) of the AVR species showed a relatively

high content of β-sheet structure as predicted for avidin and streptavidin (data not shown). Since the 'inflammation avidin' has been detected based on its ability to bind biotin and react with a polyclonal avidin antibody (Tuohimaa et al., 1989), it seems possible that these methods are unable to distinguish avidin from putative avidin-related proteins. Altogether, this suggests that the 'inflammation avidin' could be composed of avidin and/or (an) AVRs.

In conclusion, this study indicates that besides the avidin gene, there are five structurally related genes avr1–avr5 in chicken. Avidin is induced in a progesterone-specific manner and in connection with inflammation. Since avr2 and avr3 transcripts have been detected (Kunnas et al., 1993), it will be interesting to compare how the transcription of avidin and avidin-related genes is regulated. Thus the avidin gene family offers a promising model for the study of the complex mechanisms underlying the reproductive and defense systems.

Foundation (New York, NY. USA), the Academy of Finland and the Finnish Cultural Foundation.

REFERENCES

Ansorge, W., Sproat, B., Stegemann, J., Schwager, C. & Zenke, M. (1987) Automated DNA sequencing: ultrasensitive detection of fluorescent bands during electrophoresis. *Nucleic Acids Res.* 15, 4593—4602.

Benson, S. A. & Taylor, R. K. (1984) A rapid small-scale procedure for isolation of phage λ DNA. *Biotechniques* 2, 126—127.

Benton, W. D. & Davis, R. W. (1977) Screening λgt recombinant clones by hybridization to single plaques *in situ*. *Science* 196, 180—182.

Birnboim, H. C. & Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7, 1513—1523.

Chou, P. Y. & Fasman, G. D. (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* 47, 45—148.

DeLange, R. J. & Huang, T.-S. (1971) Egg-white avidin. III. Sequence of the 78-residue middle cyanogen bromide peptide. Complete amino acid sequence of the protein subunit. *J. Biol. Chem.* 246, 698—709.

Devereux, J., Haeberli, P. & Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387—395.

Elo, H. A. (1980) Occurrence of avidin-like biotin-binding capacity in various vertebrate tissues and its induction by tissue injury. *Comp. Biochem. Physiol. B Comp. Biochem.* 67, 221—224.

Elo, H., Tuohimaa, P. & Jänne, O. (1975) Cumulative superinduction of avidin in the chick oviduct by tissue damage and actinomycin D. *Mol. Cell. Endocrinol.* 2, 203—211.

Elo, H. A., Räisänen, S. & Tuohimaa, P. J. (1980) Induction of an antimicrobial biotin-binding egg white protein (avidin) in chick tissues in septic *Escherichia coli* infection. *Experientia (Basel)* 36, 312—313.

Gitlin, G., Bayer, E. A. & Wilchek, M. (1987) Studies on the biotin-binding site of avidin. Lysine residues involved in the active site. *Biochem. J.* 242, 923—926.

Gitlin, G., Bayer, E. A. & Wilchek, M. (1988a) Studies on the biotin-binding site of avidin. Tryptophan residues involved in the active site. *Biochem. J.* 250, 291—294.

Gitlin, G., Bayer, E. A. & Wilchek, M. (1988b) Studies on the biotin-binding site of streptavidin. *Biochem. J.* 256, 279—282.

Gitlin, G., Bayer, E. A. & Wilchek, M. (1990) Studies on the biotin binding sites of avidin and streptavidin. Tyrosine residues are involved in the binding site. *Biochem. J.* 269, 527—530.

Gope, M. L., Keinänen, R. A., Kristo, P. A., Conneely, O. M., Beattie, W. G., Zarucki-Schultz, T., O'Malley, B. W. & Kulomaa, M. S. (1987) Molecular cloning of the chicken avidin cDNA. *Nucleic Acids Res.* 15, 3595—3606.

Green, N. M. (1975) Avidin. *Adv. Protein Chem.* 29, 85—133.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557—580.

Heinonen, P. K. & Tuohimaa, P. (1978) The effects of 6-hydroxy-dopamine, atropin and propranolol on avidin synthesis by the chick oviduct. *Med. Biol. (Helsinki)* 56, 161—162.

Heinonen, P. K., Maisi-Ketelä, T. & Tuohimaa, P. (1978) Induction of avidin in the chick oviduct by tissue damage and prostaglandins. *J. Reprod. Fertil.* 52, 159—162.

Hertz, R. & Sebrell, W. H. (1942) Occurrence of avidin in the oviduct and secretions of the genital tract of several species. *Science* 96, 257.

Hiller, Y., Gershoni, J. M., Bayer, E. A. & Wilchek, M. (1987) Biotin binding to avidin. Oligosaccharide side chain not required for ligand association. *Biochem. J.* 248, 167—171.

Jones, P. D. & Briggs, M. H. (1962) The distribution of avidin. *Life Sci.* 11, 621—623.

Keinänen, R. A., Laukkanen, M.-L. & Kulomaa, M. S. (1988) Molecular cloning of three structurally related genes for chicken avidin. *J. Steroid Biochem.* 30, 17—21.

Klainsck, D. A., Beattie, W. G., Tsai, M.-J. & O'Malley, B. W. (1986) Molecular cloning of a steroid-regulated 108 K heat shock protein gene from hen oviduct. *Nucleic Acids Res.* 14, 10053—10067.

Korpela, J. K., Kulomaa, M. S., Elo, H. A. & Tuohimaa, P. J. (1981) Biotin-binding proteins in eggs of oviparous vertebrates. *Experientia (Basel)* 37, 1065—1066.

Korpela, J., Kulomaa, M., Tuohimaa, P. & Vaheri, A. (1982) Induction of avidin in chickens infected with the acute leukemia virus OK 10. *Int. J. Cancer* 30, 461—464.

Korpela, J., Kulomaa, M., Tuohimaa, P. & Vaheri, A. (1983) Avidin is induced in chicken embryo fibroblasts by viral transformation and cell damage. *EMBO J.* 2, 1715—1719.

Kunnas, T. A., Wallén, M. J. & Kulomaa, M. S. (1993) Induction of chicken avidin and related mRNAs after bacterial infection. *Biochim. Biophys. Acta* 1216, 441—445.

Kurzban, G. P., Gitlin, G., Bayer, E. A., Wilchek, M. & Horowitz, P. M. (1989) Shielding of tryptophan residues of avidin by binding of biotin. *Biochemistry* 28, 8537—8542.

Livnah, O., Bayer, E. A., Wilchek, M. & Sussman, J. L. (1993) Three-dimensional structures of avidin and the avidin-biotin complex. *Proc. Natl. Acad. Sci. USA* 90, 5076—5080.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Niemelä, A. O. (1986) Regulation of avidin accumulation by prostaglandins in chick oviduct culture. *J. Steroid Biochem.* 24, 709—713.

Nordback, I., Joensuu, T. & Tuohimaa, P. (1982) Effects of glucocorticoids and disodium cromoglycate on avidin production in chick tissues. *J. Endocrinol.* 92, 283—291.

O'Malley, B. W., McGuire, W. L., Kohler, P. O. & Korenman, S. G. (1969) Studies on the mechanism of steroid hormone regulation of synthesis of specific proteins. *Recent Prog. Horm. Res.* 25, 105—160.

Pugliese, L., Coda, A., Malcovati, M. & Bolognesi, M. (1993) Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 Å resolution. *J. Mol. Biol.* 231, 698—710.

Rigby, B. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113, 237—251.

Sanger, F. S., Nicklen, S. & Coulsen, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463—5467.

Senapathy, P., Shapiro, M. B. & Harris, N. L. (1990) Splice junctions, branch point sites, and exons: Sequence statistics, identification, and applications to genome project. *Methods Enzymol.* 183, 252—278.

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503—517.

Strauss, W. M. (1988) in *Current protocols in molecular biology* (Ausubel. F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds) vol. 1, pp. 2.2.1.—2.2.3., John Wiley & Sons, New York.

Tuohimaa, P., Joensuu, T., Isola, J., Keinänen, R., Kunnas, T., Niemelä, A., Pekki, A., Wallén, M., Ylikomi, T. & Kulomaa, M. (1989) Development of progestin-specific response in the chicken oviduct. *Int. J. Dev. Biol.* 33, 125—136.

Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J. & Salemme, F. R. (1989) Structural origins of high-affinity biotin binding to streptavidin. *Science* 243, 85—88.

Biotin-binding proteins in eggs of oviparous vertebrates

J.K. Korpela, M.S. Kulomaa, H.A. Elo and P.J. Tuohimaa[1]

Department of Biomedical Sciences, University of Tampere, Box 607, SF-33101 Tampere 10 (Finland), 12 February 1981

Summary. Biotin-binding was found in the egg whites and yolks of all 23 avian species studied, and in a turtle, but the amount varied considerably even in related species. There was no clear correlation in biotin-binding between egg white and yolk in various species. Antigenic determinants of avidin in different species have changed in the course of evolution as compared with those of chicken egg white avidin.

Avian egg white and the egg jelly of the frog contain a specific biotin-binding protein called avidin[2-6]. A biotin-binding protein distinct from avidin has recently been discovered in the chicken egg yolk[7,8]. In contrast to avidin, the yolk biotin-binding protein is saturated with biotin, and a special biotin-exchange assay is therefore required for its determination[7]. No comparative study of the occurrence of the egg white and yolk biotin-binding proteins has as yet been made. We therefore studied these proteins in the egg white and yolk in a number of avian and reptilian species, in fish hard roe, bull sperm and human seminal plasma.

Materials and methods. The eggs of 23 avian species, as shown in the table, were collected in southern Finland during the breeding period (April–July). 2 eggs of a turtle (*Testudo hermanni*) were utilized immediately after laying. The egg white and yolk samples were taken using separate syringes to avoid contamination, and diluted with the homogenization buffer used in the avidin assay[9]. The egg white and yolk samples, bull sperm and human seminal plasma were stored at −20°C until assayed. The hard roe of whitefish *(Coregonus albula)* or perch *(Perca fluviatilis)* were assayed fresh.

Biotin-binding in the egg whites was assayed at room temperature as previously described[9]. Egg whites of 1 species in each family were also incubated at 100°C to study the biotin saturation level and to show whether the protein has a heat stability similar to that of chicken avidin[10]. The biotin-exchange assay for egg yolks and hard roes was carried out at 65°C as described by White et al.[7], and the ^{14}C-biotin-binding reaction in other samples at room temperature (21–22°C). The avidin radioimmunoassay[9] was utilized to study in various species the presence of the antigenic determinants recognized by antiserum against chicken egg white avidin. The lipid material in the egg yolks was extracted with 1-butanol[7].

Results and discussion. Biotin-binding was found in all egg whites and yolks studied (table) including *Larus argenta-*tus[5]. The biotin-binding activity varied considerably in the egg whites of different avian species even within the same family. In the major avian families, elevated temperature (100°C) did not increase the ^{14}C-biotin-binding in the egg

Biotin-binding activities in the egg white and yolk in various avian and reptilian species

Species[a]	Number of eggs	^{14}C-biotin bound (cpm × 10³/g)[b]	
		Egg white	Egg yolk
Reptiles			
Chelonia			
Testudinidae			
Testudo hermanni	2	11.4	35.2
Birds			
Gaviiformes			
Gaviidae			
Gavia arctica	1	18	460
Anseriformes			
Anatidae			
Anas platyrhynchos[c]	9	33±4	59±2
Anas platyrhynchos[d]	7	180±29	84±4
Somateria mollissima	8	53±7	48±7
Melanitta fusca	5	218±29	209±4
Bucephala clangula	15	139±9	134±11
Galliformes			
Gallinae			
Gallus domesticus	10	290±13	88±2
Phasinidae			
Coturnix coturnix	9	513±64	57±7
Gruiformes			
Rallidae			
Fulica atra	4	372±44	130±15
Charadriiformes			
Haematopodidae			
Haematopus ostralegus	1	77	59
Stercorariidae			
Stercorarius parasiticus	1	165	–
Laridae			
Larus ridibundus	15	37±4	132±15
Larus argentatus	3	20±11	123±11
Larus canus	2	57	183
Sternidae			
Sterna hirundo	5	158±18	134±18
Columbiformes			
Columbidae			
Columba livia	2	57	183
Passeriformes			
Turdidae			
Phoenicurus phoenicurus	3	73±1	156±13
Turdus pilaris	4	26±3	97±9
Muscicapidae			
Muscicapa striata	3	9±0	136±7
Ficedula hypoleuca	6	46±7	48±4
Paridae			
Parus major	9	317±20	233±11
Corvidae			
Pica pica	2	22	143
Corvus corone	8	106±15	154±22

[a]Orders and families are also indicated. [b]The means ± SEM are given. [c]Domestic form. [d]Wild form.
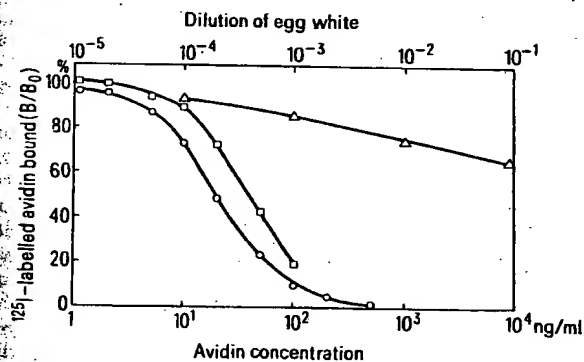
Displacement of ^{125}I-avidin with chicken and quail egg white. Serial dilutions of the chicken (□) and quail (△) egg white were assayed by the radioimmunoassay for chicken avidin. Each point represents the mean of 4 determinations. Purified chicken avidin was used to obtain an avidin standard curve (○).

white (not shown), which suggests that avidin is essentially biotin-free. On the other hand, heating decreased biotin-binding values in some avian species, suggesting a smaller stability of their avidin to heat than that in the chicken.

The quail egg white showed an incomplete cross-reaction in the radioimmunoassay for chicken avidin (fig.), while the egg whites of other avian species could not prevent ^{125}I-labelled avidin from binding to antiserum. This result indicates differences in the antigenic determinants of avidin molecules as compared to chicken avidin.

The biotin-binding protein found in the chicken egg yolk[7,11] was demonstrated here to be a common constituent in the egg yolk of various avian species (table). This protein is distinct from avidin, since it is denaturated at 100 °C[7,10]. The biotin-binding activities in the egg yolks also varied considerably from species to species. No clear correlation was found between the biotin-binding activities in the egg white and yolk in different species. The lipid-free yolk material in all avian species studied did not show any cross-reaction in the avidin radioimmunoassay.

The egg white and yolk of the turtle also showed biotin-binding activity (table), but no cross-reaction in the avidin radioimmunoassay. No biotin-binding activity was found in the hard roe of the fishes, bull sperm or human seminal plasma. Jones and Briggs[5] discovered a low biotin-binding activity in fresh bull sperm. This discrepancy in results might originate in the microbiological avidin assay they used, since the growth of microbes could be inhibited by any growth inhibitor present in the bull sperm.

It has been proposed that the biotin-binding proteins might be widely distributed in the animal kingdom and play some fundamental role in the physiology of reproduction[3-5]. An antimicrobial[6,12] effect for avidin, and a biotin-transporting role[11] for yolk biotin-binding protein, have been suggested as their functions. Fishes and mammals[13-16] so far studied did not contain any biotin-binding protein similar to that found in egg whites and yolks. It seems possible that special biotin-binding proteins have evolved for reproductive purposes in amphibian, reptilian and avian eggs.

1 We thank Mr Jukka Peltonen and Mr Antti Karlin for the collection of the avian eggs with permission obtained from the Ministry of Agriculture, and Mr Reino Saarinen for the turtle eggs. The authors are indebted to Mrs Outi Kurronen, Miss Riitta Mero and Miss Tiina-Maija Mattila for technical assistance. This work was supported by the Ford Foundation Grants No. 760-0526 and No. 790-0665.

2 R. E. Eakin, E. E. Snell and R. J. Williams, J. biol. Chem. 140, 535 (1941).

3 R. Hertz and W. H. Sebrell, Science 96, 257 (1942).

4 R. E. Feeney, J. S. Anderson, P. R. Azari, N. Bennett and M. B. Rhodes, J. biol. Chem. 235, 2307 (1960).

5 P. D. Jones and M. H. Briggs, Life Sci. 11, 621 (1962).

6 N. M. Green, Adv. Protein Chem. 29, 85 (1975).

7 H. B. White III, B. A. Dennison, M. A. Della Fera, C. J. Whitney, J. C. McGuire, H. W. Meslar and P. H. Sammelwitz, Biochem. J. 157, 395 (1976).

8 H. W. Meslar, S. A. Camper and H. B. White III, J. biol. Chem. 253, 6979 (1978).

9 M. S. Kulomaa, H. A. Elo and P. J. Tuohimaa, Biochem. J. 175, 685 (1978).

10 M. S. Kulomaa, H. A. Elo, A. O. Niemelä and P. J. Tuohimaa. Biochim. biophys. Acta 670, in press (1981).

11 R. D. Mandella, H. W. Meslar and H. B. White III, Biochem. J. 175, 629 (1978).

12 H. A. Elo, S. Räisänen and P. J. Tuohimaa, Experientia 36, 312 (1980).

13 R. Hertz, Physiol. Rev. 26, 479 (1946).

14 H. A. Elo, M. S. Kulomaa and P. J. Tuohimaa, Comp. Biochem. Physiol. 62B, 237 (1979).

15 H. A. Elo, Comp. Biochem. Physiol. 67B, 221 (1980).

16 P. Tuohimaa, M. Kulomaa, A. Niemelä, T. Torkkeli, O. Jänne and S. J. Segal, Proc. natl Acad. Sci. USA, submitted for publication.

A low molecular weight tracer molecule for immunocytochemistry. Identification of cytoplasmic actin

R. Tiggemann and M. V. Govindan

Faculty of Biology, University of Konstanz, D–7750 Konstanz (Federal Republic of Germany), and German Cancer Research Centre, Im Neuenheimer Feld 280, D–6900 Heidelberg (Federal Republic of Germany), 11 February 1981

Summary. Anti actin Fab-fragments were tagged to a small electron dense tracer molecule; ferrocene monocarboxylic acid (230 daltons). The conjugate stains actin filaments, which were found mainly in the core of microvilli.

Many technical efforts have been made to visualize antigenic structures immunocytochemically. The main obstacle has been that the methods depend upon very large tracer molecules such as ferritin[1]. The immunoperoxidase method[2] does not eliminate this problem either, as the enzyme-antibody complex is still too large in diameter to pass through the cell membrane. Attempts were also made to allow large molecules to penetrate the plasma membrane with membrane disrupting agents[3,4], or with enzymatic digestion of certain membrane components[5]; however, these manipulations resulted in the destruction of the cell shape. Thus most techniques are still far from being established for immunocytochemistry. We here present a more suitable staining procedure, which helps to avoid most of the difficulties mentioned above. The very small Fab-fragment – ferrocene carboxylic acid (FMCA) complex identifies intracellular antigens without background effects. The procedure is easy to handle, direct and does not result in the destruction of cellular ultrastructure.

Actin was isolated from Ehrlich mouse ascites tumour (MAT) cells according to Lazarides and Weber[6], purified by polymerization- depolymerization cycles[7] and by DNase-I affinity chromatography[8]. Actin was injected s.c. into male New Zealand rabbits in the presence of complete Freund's adjuvant (protein concentration: 1.5 mg/ml). This was repeated on days 8 and 40 after the 1st inoculation. The IgG fraction was isolated from the serum and purified by DEAE-52 ion exchange chromatography[9]. Fab-fragments were prepared according to Porter[10] and labeled with the iron-containing sandwich molecule FMCA, using a water soluble carbodiimide[11]. Fab-fragments (5 mg protein), FMCA (5 mg) and 1-ethyl-3 (3-dimethylaminopropyl)-carbodiimide (10 mg) were dissolved in 2.5 ml 10 mM sodium phosphate and gently stirred at 4 °C over night. The

Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system

Anke Krebber, Susanne Bornhauser, Jörg Burmester, Annemarie Honegger, Jörg Willuda, Hans Rudolf Bosshard, Andreas Plückthun *

Biochemisches Institut der Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

Abstract

A prerequisite for the use of recombinant antibody technologies starting from hybridomas or immune repertoires is the reliable cloning of functional immunoglobulin genes. For this purpose, a standard phage display system was optimized for robustness, vector stability, tight control of scFv-ΔgeneIII expression, primer usage for PCR amplification of variable region genes, scFv assembly strategy and subsequent directional cloning using a single rare cutting restriction enzyme. This integrated cloning, screening and selection system allowed us to rapidly obtain antigen binding scFvs derived from spleen-cell repertoires of mice immunized with ampicillin as well as from all hybridoma cell lines tested to date. As representative examples, cloning of monoclonal antibodies against a his tag, leucine zippers, the tumor marker EGP-2 and the insecticide DDT is presented. Several hybridomas whose genes could not be cloned in previous experimental setups, but were successfully obtained with the present system, expressed high amounts of aberrant heavy and light chain mRNAs, which were amplified by PCR and greatly exceeded the amount of binding antibody sequences. These contaminating variable region genes were successfully eliminated by employing the optimized phage display system, thus avoiding time consuming sequencing of non-binding scFv genes. To maximize soluble expression of functional scFvs subsequent to cloning, a compatible vector series to simplify modification, detection, multimerization and rapid purification of recombinant antibody fragments was constructed.

Keywords: Phage display; Single-chain Fv; Monoclonal antibody; Antibody library

---

1. Introduction

Molecular cloning and sequencing of antibody variable domains forms the basis of antibody modelling (Rees et al., 1994), antibody engineering (Plückthun, 1994; Nilsson, 1995) and experimental structure determination by NMR (Freund et al., 1994) or X-ray crystallography at high resolution (Ostermeier et al., 1995). Moreover, once the variable region genes have been cloned, the antibody domains can be further engineered in a multitude of ways to produce antibody variants with lower immunogenicity (Güssow and Seemann, 1991), higher affinity (Marks et al., 1992; Riechmann and Weill, 1993; Deng et al., 1994), altered antigenic specificity (Ohlin et al., 1996), or enhanced stability (Glockshuber et al., 1990; Reiter et al., 1994). Furthermore, genetic fusions of scFv fragments to effector proteins and toxins are powerful tools in the fields of medicine and diagnostics (Huston et al., 1993).

In all application areas, the demand for efficient generation of functional antibody fragments increases continuously. Although large prefabricated antibody libraries are gradually becoming a source of recombinant antibody fragments that cover a wide range of useful affinities (Vaughan et al., 1996), it may still be necessary to use the diversity of the immune system to create the most extensive panel of different antibodies against a given target possible. Furthermore, it is often of great interest and importance to clone V_H and V_L domains of the natural antibody response to a given antigen. In cases in which a large amount of experimental or clinical data is available on a given monoclonal antibody (mAb), it is frequently useful to base new constructs on this work and to determine its specific sequence and binding mode. Cloning and sequencing retains and immortalizes the unique and extensively characterized specificity of mAbs, which can be crucial for the rescue of unstable hybridoma cell lines.

One major problem in rapidly and simply obtaining sequence information about mAbs stems from the occurrence of aberrant mRNAs which are transcribed from rearranged, but non-functional, heavy and light chain genes in the hybridoma (Cabilly and Riggs, 1985; Strohal et al., 1987; Carroll et al., 1988; Kaluza et al., 1992; Kütemeier et al., 1992; Nicholls

et al., 1993; Duan and Pomerantz, 1994; Yamanaka et al., 1995; Ostermeier and Michel, 1996). These non-productive chains are frequently preferentially amplified over the productive ones by sets of primers specific for the variable regions of antibody genes. The aberrant chains may greatly dilute the desired antibody sequences, which are the only ones binding the antigen in a pool of non-productive antibody-like sequences. Several attempts have been reported to overcome this problem, such as ribozyme cleavage of a known aberrant κ chain sequence (Duan and Pomerantz, 1994), treatment of aberrant mRNA/DNA hybrids with RNAseH (Ostermeier and Michel, 1996), or functional screening for full length scFv products in an in vitro transcription/translation system (Nicholls et al., 1993). Each of these methods is time consuming, depends on prior sequence information of the contaminating gene and fails to enrich binding molecules by selection procedures. Since antibody genes are usually amplified by PCR using degenerate sets of primers, mismatches and PCR errors will lead to point mutations or out-of-frame clones, which can also contribute to a background of non-functional scFv molecules. Therefore, it is absolutely vital, but often neglected (Miller et al., 1995; Kwak et al., 1996), that the binding specificity of the recombinant antibody sequence is demonstrated to be comparable with the binding characteristics of the parental monoclonal antibody, even when the deduced antibody sequence seems reasonable.

The inherent advantage of phage display is its direct link of DNA sequence to protein function (McCafferty et al., 1990; Winter et al., 1994). Thus, single clones can be rapidly screened for antigen binding and, even more importantly, selected from pools in the same experimental setup. This obviates the use of sequence specific methods to eliminate undesired sequences and leads to a more generally applicable procedure for hybridoma cloning.

However, phage display suffers from the fact that non-productive, aberrant chains are often very well expressed and non-toxic to the bacterial cell, whereas cells expressing functional scFv-geneIII fusions have a growth disadvantage and are selected against. The scFv-geneIII fusion protein can cause vector instability, creating deletions in the antibody fusion genes as occasionally observed (Courtney et al., 1995; Dziegiel et al., 1995; A. Krebber, unpublished obser-

vations; footnote 1). Thus, it is highly recommended to use a regulatable vector system allowing tight product suppression during all propagation steps as well as controlled expression of low amounts of scFv-geneIII fusion protein for phage display. Since a variety of serious technical problems concerning hybridoma cloning and enrichment of binding antibody fragments from phage display libraries have been reported [1], we have developed the reengineered phage display system described in this work. In order to provide a robust and straightforward methodology which ensures fast and reliable cloning, not only of hybridomas but also of larger antibody libraries, each step in the process was optimized. To illustrate the utility of our improved phage display system we report in detail several case studies of successfully cloned scFvs derived from monoclonal antibodies as well as enrichment of binding scFv sequences from cloned B cell repertoires.

2. Materials and methods

2.1. Isotyping

Isotypes of the mAbs were determined using the IsoStrip mouse monoclonal antibody isotyping kit (Boehringer Mannheim).

2.2. Preparation of mRNA

mRNA was extracted from $1–5 \times 10^6$ hybridoma or spleen cells using the QuickPrep mRNA purification kit from Pharmacia. In the case of hybridoma cell lines 13AD and 42PF total RNA was isolated essentially as described by Berger and Chirgwin (1989).

2.3. First strand cDNA synthesis

About 1 μg mRNA or 5 μg total RNA was reverse transcribed in a reaction volume of 33 μl using random hexamer primers according to the

manufacturer's protocol (first strand cDNA synthesis kit (Pharmacia)).

2.4. PCR amplification of V_L and V_H

Various DNA polymerases (Taq (Perkin Elmer, Gibco), Pwo (Boehringer Mannheim), Pfu (Stratagene), Vent (New England Biolabs)) were successfully used for separate amplifications of V_L and V_H. For amplification of V_L from hybridomas either λ or κ primers were chosen according to the isotype. PCR reactions were performed in 50–100 μl volumes, containing 2–5 μl of cDNA reaction, 2 μM of LB and LF primer mixes (Table 1, Fig. 1B) for amplification of V_L or 2 μM of HB and HF primer mixes (Table 1, Fig. 1B) for amplification of V_H, 200 μM dNTPs, an optimized Mg^{2+} concentration (2–6 mM) and reaction buffer supplied by the manufacturers. After 3 min denaturation at 92°C, 2 U of DNA polymerase were added, followed by 7 cycles of 1 min at 92°C, 30 s at 63°C, 50 s at 58°C, 1 min at 72°C, and 23 cycles of 1 min at 92°C, 30 s at 63°C, 1 min at 72°C. One tenth of each PCR reaction was analyzed by agarose gel electrophoresis (Fig. 2).

2.5. Assembly PCR

The full length PCR products of V_L and V_H were purified by preparative agarose gel electrophoresis in combination with the QIAEX (Qiagen) or Jetsorb (Genomed) DNA extraction kit. Approximately 10 ng of each V_L and V_H DNA were combined by SOE-PCR (Fig. 1C; Ge et al., 1995). An initial denaturation step (3 min, 92°C) was followed by 2 cycles of 1 min at 92°C, 30 s at 63°C, 50 s at 58°C, 1 min at 72°C in the absence of primers. After adding the outer primers scback and scfor (Table 1, Fig. 1C; each 1 μM), 5 cycles of 1 min at 92°C, 30 s at 63°C, 50 s at 58°C, 1 min at 72°C, and 23 cycles of 1 min at 92°C, 30 s at 63°C, 1 min at 72°C were performed. One tenth of each PCR reaction was analyzed by agarose gel electrophoresis (Fig. 2).

2.6. SfiI digest and cloning of scFv fragments into pAK100

The gel-purified scFv fragment and the phage display vector pAK100 were both digested with SfiI

---

[1] For typical examples, see the internet discussion forums,
http://www.bio.net/hypermail/METHDS-REAGNTS,
http://www.bio.net/hypermail/MOLECULAR-REPERTOIRES.

for 3 to 4 h at 50°C (Fig. 1D). After purification, the scFv fragment was ligated into the vector (molar ratio vector to insert 1.5 : 1) and transformed into *E. coli* XL1-Blue (Stratagene). For library construction, ligation mixtures precipitated with n-butanol (Thomas, 1994) were electroporated into XL1 Blue (Dower et al., 1988; yield approximately 5×10^7 clones per μg SfiI-cut insert DNA). After plating on NE medium (non expression medium: 2 × YT containing 1% glucose and 25 μg/ml chloramphenicol)

Table 1

Listing of the primers used for assembling mouse scFv fragments in the orientation VL-$(G_4S)_4$-VH, which are compatible with the pAK vector system presented in Fig. 1 and Fig. 4

Primer VL back:

```
             5'                        FLAG   VL3'
scback  ttactcgcggcccagccggccatggcggactacaaaG
```

	5' FLAG VL 3'	d	µl Mix
LB1	gccatggcggactacaaaGAYATCCAGCTGACTCAGCC	2	1
LB2	gccatggcggactacaaaGAYATTGTTCTCWCCCAGTC	4	2
LB3	gccatggcggactacaaaGAYATTGTGMTMACTCAGTC	8	5
LB4	gccatggcggactacaaaGAYATTGTGYTRACACAGTC	8	3.5
LB5	gccatggcggactacaaaGAYATTGTRATGACMCAGTC	8	4
LB6	gccatggcggactacaaaGAYATTMAGATRAMCCAGTC	16	7
LB7	gccatggcggactacaaaGAYATTCAGATGAYDCAGTC	12	6
LB8	gccatggcggactacaaaGAYATYCAGATGACACAGAC	4	1.5
LB9	gccatggcggactacaaaGAYATTGTTCTCAWCCAGTC	4	2
LB10	gccatggcggactacaaaGAYATTGWGCTSACCCAATC	8	3.5
LB11	gccatggcggactacaaaGAYATTSTRATGACCCARTC	16	8
LB12	gccatggcggactacaaaGAYRTTKTGATGACCCARAC	16	8
LB13	gccatggcggactacaaaGAYATTGTGATGACBCAGKC	12	6
LB14	gccatggcggactacaaaGAYATTGTGATAACYCAGGA	4	2
LB15	gccatggcggactacaaaGAYATTGTGATGACCCAGWT	4	2
LB16	gccatggcggactacaaaGAYATTGTGATGACACAACC	2	1
LB17	gccatggcggactacaaaGAYATTTTGCTGACTCAGTC	2	1
LBλ	gccatggcggactacaaaGATGCTGTTGTGACTCAGGAATC	1	1

Primer VL for:

	5' (Gly4Ser)3-linker VL 3'		
LF1	ggagccgccgccgcc(agaaccaccaccacc)2ACGTTTGATTTCCAGCTTGG	1	1
LF2	ggagccgccgccgcc(agaaccaccaccacc)2ACGTTTTATTTCCAGCTTGG	1	1
LF4	ggagccgccgccgcc(agaaccaccaccacc)2ACGTTTTATTTCCAACTTTG	1	1
LF5	ggagccgccgccgcc(agaaccaccaccacc)2ACGTTTCAGCTCCAGCTTGG	1	1
LFλ	ggagccgccgccgcc(agaaccaccaccacc)2ACCTAGGACAGTCAGTTTGG	1	0.25

Primer VH back:

	5' (Gly4Ser)2-linker BamHI VH 3'	d	µl Mix
HB1	ggcggcggcggctccggtggtggtggatccGAKGTRMAGCTTCAGGAGTC	8	4
HB2	ggcggcggcggctccggtggtggtggatccGAGGTBCAGCTBCAGCAGTC	9	4
HB3	ggcggcggcggctccggtggtggtggatccCAGGTGCAGCTGAAGSASTC	4	3
HB4	ggcggcggcggctccggtggtggtggatccGAGGTCCARCTGCAACARTC	4	4
HB5	ggcggcggcggctccggtggtggtggatccCAGGTYCAGCTBCAGCARTC	12	7
HB6	ggcggcggcggctccggtggtggtggatccCAGGTYCARCTGCAGCAGTC	4	2
HB7	ggcggcggcggctccggtggtggtggatccCAGGTCCACGTGAAGCAGTC	1	1
HB8	ggcggcggcggctccggtggtggtggatccGAGGTGAASSTGGTGGAATC	4	2
HB9	ggcggcggcggctccggtggtggtggatccGAVGTGAWGYTGGTGGAGTC	12	5
HB10	ggcggcggcggctccggtggtggtggatccGAGGTGCAGSKGGTGGAGTC	4	2
HB11	ggcggcggcggctccggtggtggtggatccGAKGTGCAMCTGGTGGAGTC	4	2
HB12	ggcggcggcggctccggtggtggtggatccGAGGTGAAGCTGATGGARTC	2	2
HB13	ggcggcggcggctccggtggtggtggatccGAGGTGCARCTTGTTGAGTC	2	1
HB14	ggcggcggcggctccggtggtggtggatccGARGTRAAGCTTCTCGAGTC	4	2
HB15	ggcggcggcggctccggtggtggtggatccGAAGTGAARSTTGAGGAGTC	4	2
HB16	ggcggcggcggctccggtggtggtggatccCAGGTTACTCTRAAAGWGTSTG	8	5
HB17	ggcggcggcggctccggtggtggtggatccCAGGTCCAACTVCAGCARCC	6	3.5
HB18	ggcggcggcggctccggtggtggtggatccGATGTGAACTTGGAAGTGTC	1	0.7
HB19	ggcggcggcggctccggtggtggtggatccGAGGTGAAGGTCATCGAGTC	1	0.7

Primer VH for:

	5'EcoRI 3'		
scfor	ggaattcggccccgag		

	5'EcoRI VH 3'		
HF1	ggaattcggccccgaggccCGAGGAAACGGTGACCGTGGT	1	1
HF2	ggaattcggccccgaggccCGAGGAGACTGTGAGAGTGGT	1	1
HF3	ggaattcggccccgaggccCGCAGAGACAGTGACCAGAGT	1	1
HF4	ggaattcggccccgaggccCGAGGAGACGGTGACTGAGGT	1	1

in 530 cm^2 dishes (Nunc) and overnight incubation at RT, the colonies were scraped off the plates into 8 ml 2 × YT (Sambrook et al., 1989) and subsequently stored at −80°C after addition of 10% glycerol.

2.7. Rescue of scFv displaying phages

To rescue scFv displaying phages, 50 ml NE medium was inoculated with approximately 10^9 cells from the glycerol library stock. The culture was then shaken at 37°C. At $OD_{550} = 0.5$, 10^{11} cfu helper phage VCS (Stratagene) and 25 μl 1 M IPTG solution were added. After 15 min incubation at 37°C without agitation, the culture was diluted in 100 ml LE medium (low expression medium: 2 × YT containing 1% glucose, 25 μg/ml chloramphenicol and 0.5 mM IPTG). The culture was then shaken for 10 h at 26°C for phage production. 2 h after infection 30 μg/ml kanamycin was added. Phage particles were purified and concentrated 100-fold by two PEG/NaCl precipitations (Sambrook et al., 1989), resuspended in PBS and stored at 4°C. After overnight culture typically a phage titer in the range of 10^{11} cfu/ml was observed.

2.8. Selection of antigen binders by panning

For selection, immunotubes (Nunc, Maxisorp) were coated overnight at 4°C with 4 ml of 10–100 μg/ml antigen solution (for the anti-ampicillin libraries: 100 μg/ml transferrin-EMCS-ampicillin in

PBSU (1 : 1 mixture of PBS (20 mM NaP_i, 150 mM NaCl, pH 7.2) and urea-NaP_i buffer (8 mM urea, 50 mM NaP_i, pH 7.0)); for anti-leucine zipper mAb 13AD: 10 μg/ml biotin-LZ_{ss}/streptavidin complex (Leder et al., 1995) in 34.8 mM $NaHCO_3$, 15 mM Na_2CO_3, 0.02% NaN_3, pH 9.6; for the anti-EGP-2 mAb MOC31: 100 μg/ml EGP-2 in PBS). After blocking with 4% dried skimmed milk powder in PBS for 2 h at room temperature (RT), 10^{11} phagemid particles in 4 ml PBS containing 2% milk were added and incubated for 2 h with rocking at RT. Tubes were then washed 15 times with PBS/0.1% Tween and 15 times with PBS. Bound phages were eluted from the tube with soluble antigen (1 ml 10 mM ampicillin in PBS; anti-ampicillin libraries) or 800 μl 0.1 M glycine/HCl pH 2.2 for 10 min. The latter solution was neutralized with 48 μl 2 M Tris and the phages (typically 10^4 –10^6 cfu/ml) were used for reinfection (30 min at 37°C without agitation) of E. coli XL1-Blue cells in NE medium (OD_{550} = 0.5–0.8). In the case of elution with ampicillin, the solution was treated with 2 units of β-lactamase for 15 min before reinfection. This sublibrary was rescued as described above and subjected to further panning rounds or binding analysis by phage ELISA.

2.9. Phage ELISA

Single colonies were grown separately in 2 ml NE medium at 37°C. After reaching $OD_{550} = 0.3$, 1 ml

---

Table 1 (continued). In this nomenclature, 'back' refers to 'toward the 3′ end of the antibody gene' and 'for' to 'toward the 5′ end of the antibody gene'. The sequences are given using the IUPAC nomenclature of mixed bases (shown in underlined capital letters, R = A or G; Y = C or T; M = A or C; K = G or T; S = C or G; W = A or T; H = A or C or T; B = C or G or T; V = A or C or G; D = A or G or T), with a column listing the d-fold degeneration encoded in each primer and the μl to be used to set up the PCR mix. The LB1-LB17 series encodes a stretch of 20 bases hybridizing to the mature mouse antibody κ sequences (in capital letters). Underlined is the preceding sequence which encodes the shortened FLAG sequence (Knappik and Plückthun, 1994). Since the FLAG tag uses the fixed N-terminal aspartate of the mature antibody (encoded by GAY), only three additional amino acids are necessary. The FLAG codons are then preceded by the codons specifying the end of the pelB signal sequence. The LBλ primer for mouse lambda chains is constructed in an analogous manner (the N-terminal glutamate of the mature mouse λ sequence is replaced by aspartate (encoded by GAT) to generate a FLAG tag). The VLfor primer sequences are complementary to the J elements of κ or λ chains (capital letters) and encode three repeats of the Gly$_4$Ser sequence, the terminal one (bold) of which has a different codon usage so that incorrect overlaps during the PCR assembly reaction are minimized. The VHback primers encode the other part of the linker as well as a BamHI recognition site (underlined), and the overlap with VLfor in the sequence shown in bold. The 20 bases given in capital letters hybridize with the mature mouse V$_H$ sequences. The last 19 nt at the 3′ end of the VHfor primers hybridize with the J$_H$ region. The first nt shown in capital letters will introduce a silent mutation at the end of V$_H$ in order to code for the first nt of the second SfiI recognition site (underlined). The final assembly of the scFv gene by SOE-PCR is carried out with the scback and scfor primer set. The outer primer scback encodes the first SfiI site (underlined). All primers contain a phosphorothioate at the 3′ end, with the exception of scfor, to avoid potential interference with SfiI digestion.

NE medium complemented with 5×10^9 cfu VCS helper phage (Stratagene) and 1.5 mM IPTG was added. The cultures were allowed to produce scFv displaying phages during overnight incubation at 24°C. Phages from 1.6 ml culture supernatants were PEG precipitated and dissolved in 200 μl PBS. For anti-leucine zipper hybridomas 13AD and 42PF, scFv displaying phages were produced and concentrated 100-fold by PEG precipitation as described for the phage panning experiments. 10–100 μg/ml antigen (for the anti-his tag hybridoma 3D5: 100 μg/ml

his-tagged citrate synthase in PBS; for the anti-leucine zipper hybridoma 42PF: 10 μg/ml BSA-LZ(7P14P) in 34.8 mM NaHCO$_3$, 15 mM Na$_2$CO$_3$, 0.02% NaN$_3$, pH 9.6; for the anti-DDT hybridoma 3D7: 70 μg/ml β-alaninamide-DDT conjugated to lysozyme; for the anti-ampicillin library and the hybridomas 13AD and MOC31: coating solutions as described for the panning procedure above) was coated onto NUNC plastic ELISA plates by overnight incubation at 4°C. 50 μl phage solution per well was mixed with 50 μl 4% dried skimmed milk powder in
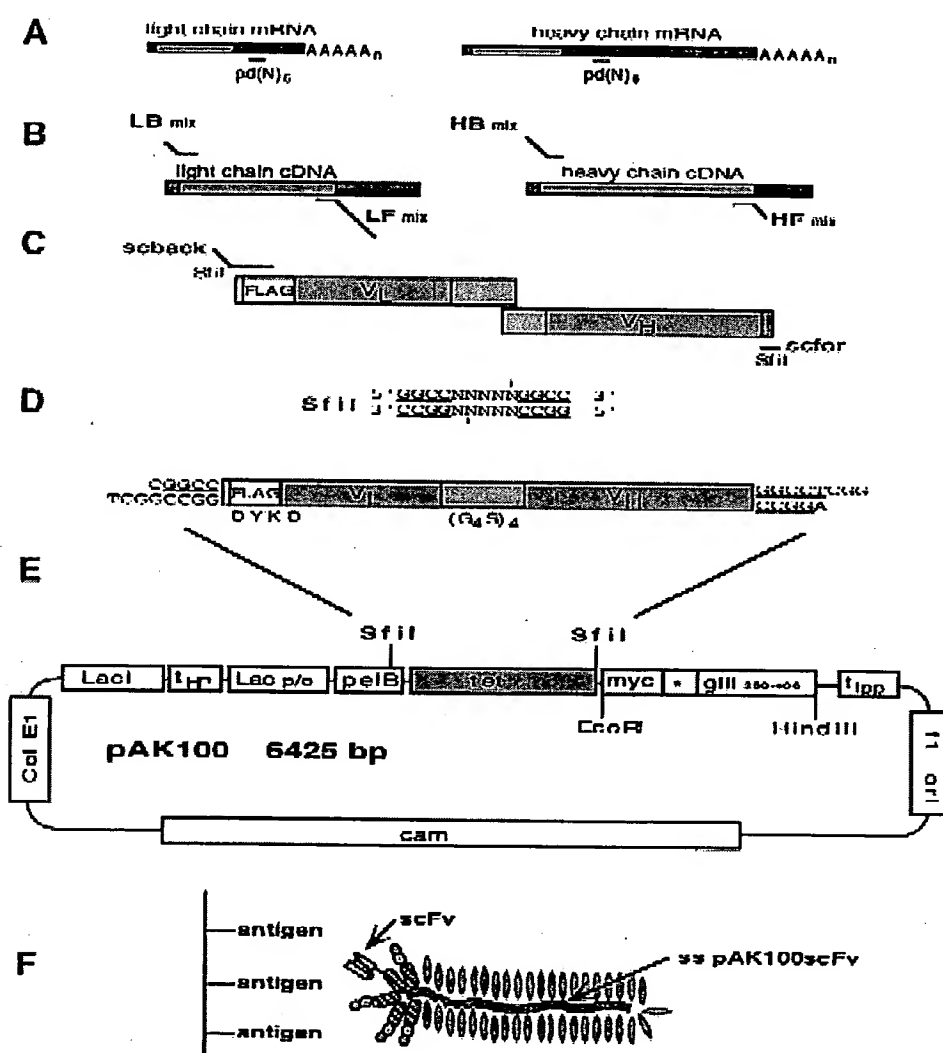
PBS in the presence or absence of competing soluble antigen and incubated at RT for 10 min. After washing and blocking (blocking buffer: 4% dried skimmed milk powder in PBS) of the ELISA wells, the phage solution was added and incubated for 1 h at RT. After washing, 100 μl of 1/5000 diluted HRP/anti M13-conjugate (Pharmacia) in blocking buffer was added and incubated for 1 h at RT. For detection, 100 μl soluble BM blue POD-substrate (Boehringer Mannheim) was used.

2.10. Soluble expression of scFv fragments in pAK300 and pAK400

For soluble expression, 20 ml expression medium (2 × YT containing 25 μg/ml chloramphenicol) was inoculated with 200 μl of preculture (JM83 harboring the expression plasmid for the anti-ampicillin scFv antibody aL2, pAK300scFvaL2 or pAK400scFvaL2, respectively; grown overnight at RT in NE medium) and incubated in a shaking waterbath at 24°C. Expression was induced at OD_{600} = 0.5 with 1 mM IPTG and allowed to proceed for 4 h at 24°C. To monitor total scFv production, an aliquot of the culture was diluted in PBS to an OD_{600} = 1 and mixed with 5 × reducing SDS-PAGE sample buffer (Sambrook et al., 1989). The rest of the culture was adjusted to OD_{600} = 5 by centrifuga-

tion and resuspension in 2 ml PBS. The cells were disrupted three times by French press and separated into soluble and insoluble fractions by centrifugation (soluble fraction = supernatant; insoluble fraction = pellet vortexed in the same buffer volume). Aliquots of both fractions were mixed with 5 × reducing SDS-PAGE sample buffer. A 12.5 μl aliquot of each sample was used for 0.1% SDS-12% PAGE and subsequent Western blot detection.

2.11. Western blot

Gels were blotted onto PVDF membranes using standard protocols. The scFv fragments were detected using the anti-FLAG mAb M1 (Kodak), followed by an anti-mouse IgG peroxidase conjugate, essentially as described in Knappik and Plückthun (1994).

2.12. Sequencing

Nucleic acid sequences were determined either by manual Sanger dideoxy sequencing (USB Sequenase kit), or by cycle sequencing (Sequi Therm Long-Read Cycle Sequencing Kit-LC, Epicentre Technologies) with fluorescent primers using a DNA sequencer (LI-COR).

Fig. 1. Scheme of amplification and cloning procedure. A: cDNA synthesis. mRNA derived from spleen cells or hybridomas and random hexamer primer $(pd(N)_6)$ or subclass specific primers (not shown) are used for cDNA synthesis. B: PCR amplification of V_L and V_H domains. The cDNA is used as the PCR template for the amplification of V_L and V_H domains by the primer mixes indicated (listed in Table 1). C: assembly SOE-PCR. V_L and V_H PCR products are first assembled into the scFv format (splicing by overlap extension) without primers and subsequently amplified by the outer primer pair scback and scfor. D: SfiI digestion of the amplified scFv fragment. The rare cutting enzyme SfiI is the only enzyme used for antibody cloning. E: ligation of SfiI digested pAK100 vector and insert. Note that directional cloning of the SfiI inserts is guaranteed because of the different SfiI sticky ends shown. In addition, self-ligation of insert or vector molecules is excluded by the asymmetry of the overhang. The phage display vector pAK100 contains a tet resistance cassette (tetA and tetR; 2101 bp) to facilitate monitoring of complete SfiI digestion by gel electrophoresis and by religating and subsequent plating on tet plates, the lacI repressor gene, a strong upstream terminator (t_{HP}), the lac promoter/operator and the pelB leader sequence, which has been modified to contain an SfiI site (for details see Fig. 3A). After ligation, the antibody fragment is fused in frame to geneIII250–406 (Fig. 3B). The in-frame fusion contains a myc-tag (Munro and Pelham, 1986) to act as a detection handle, in addition to the short 3-amino acid FLAG tag at the N-terminus (Knappik and Plückthun, 1994). The asterisk represents an amber codon. The geneIII portion starts at position 250 of the wt geneIII protein (gIIIp), thus avoiding extraordinarily long glycine linkers and, most importantly, any unpaired cysteine of gIIIp. The expression cassette is followed by a downstream terminator (t_{lpp}). The origins for phage replication and plasmid replication are as described in Ge et al. (1995). The chloramphenicol (cam) cassette is originally derived from pACYC184, but its expression strength has been modified by randomizing the promoter and selecting clones with optimal growth and selection properties (Krebber et al., 1995). F: detection and enrichment of binding scFv sequences by phage display. The scFv insert is displayed on the tip of filamentous phage whereas the genetic information encoding for the particular scFv fragment is packaged as single stranded DNA (ss pAK100scFv) in the phage interior. Panning of single-chain antibody displaying phages against the antigen allows the enrichment of functional antibody sequences.

3. Results

3.1. Design features of the improved phage display system

The reengineered phage display system and optimized methodology used in this work combines the following significantly improved features.

(i) In many cases, previously reported primer sets were too restricted to amplify either particular light or heavy chains (Table 2). Therefore, the set of mouse primers used in this study (Table 1) has been extended and optimized. It incorporates all mouse V_H, V_λ and V_κ sequences collected in the Kabat data base (Kabat et al., 1991) and combines extended primer sets described by Kettleborough et al. (1993); Ørum et al. (1993) and Zhou et al. (1994).

(ii) The V_L back primer set encodes a convenient, shortened version of the FLAG peptide, which introduces only three additional amino acids at the N-terminus of V_L. In this way, the scFv can be easily detected and purified by a commercially available mAb (Knappik and Plückthun, 1994; Ge et al., 1995; Kalinke et al., 1996).

(iii) To minimize PCR errors, polymerases with proof-reading capacity are used whenever possible (Marks et al., 1991; Yamanaka et al., 1995).

(iv) The scFv fragment is efficiently assembled by SOE-PCR (splicing by overlap extension; Horton et al., 1989) from two (Ge et al., 1995; Vaughan et al., 1996) rather than three pieces (Clackson et al., 1991; Ørum et al., 1993; recombinant phage antibody system (Pharmacia)).

(v) To avoid the occurrence of incorrect overlaps during assembly PCR, the four (Gly_4Ser) repeats in the single chain linker region are encoded by different codons (Table 1; Ge et al., 1995). In order to reduce the dimerization or aggregation tendency of scFv fragments (Desplancq et al., 1994; Huston et al., 1995), the linker between V_L and V_H is 20 amino acids in length rather than the frequently used 15 amino acids long variant.

(vi) SfiI is the only enzyme used for directional cloning of scFv fragments into the optimized phage display vector pAK100 (Fig. 1E). The use of this enzyme has a number of distinct advantages: SfiI recognizes eight bases, interrupted by five non-recognized nucleotides (5′-GGCCNNNNNGGCC-3′).

SfiI restriction sites are therefore very rare in antibody sequences, thus elimination of potentially interesting sequences by internal digestion is very unlikely. Two different sticky ends were designed to allow cloning of the scFv fragment in a directional manner. In contrast to the palindromic sticky ends, 3 bp overhangs derived from SfiI sites render impossible self-dimerization by either insert or vector. Finally, SfiI has the interesting property that it always cuts two sites at once (Wentzell et al., 1995), and therefore single-cut plasmids or inserts do not occur as intermediates. Digestion of vectors or inserts with single SfiI sites requires the binding of two different DNA molecules to the restriction enzyme and slows the turnover rate (Wentzell et al., 1995). While vectors with asymmetric SfiI sites have been described (Zelenetz, 1992; Barbas and Wagner, 1995; Yang et al., 1995), surprisingly this feature has not been used in antibody library cloning. Usually only one SfiI site, mostly in combination with NotI as a second site, is used (Hoogenboom et al., 1991; Ørum et al., 1993; Vaughan et al., 1996). Other systems employ a set of four enzymes to clone V_L and V_H independently of each other (Orlandi et al., 1989; Barbas et al., 1991; Johansen et al., 1995; Yamanaka et al., 1995). These systems thus run a significantly
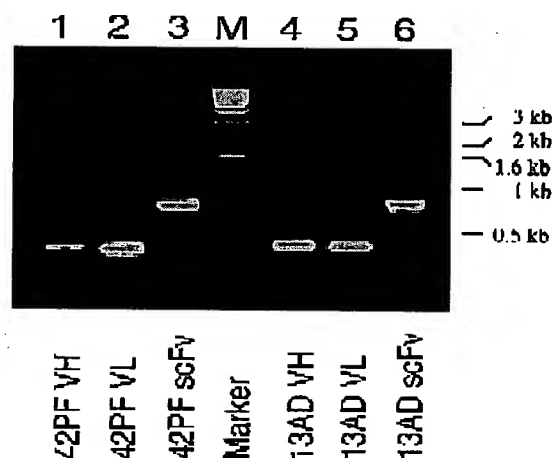
Fig. 2. DNA gel. PCR amplification of V_L and V_H from mouse hybridomas 13AD and 42PF using the primer sets listed in Table 1. Lane 1: V_H mAb 42PF; lane 2: V_L mAb 42PF; lane 3: assembled scFv 42PF; M: 1 kb marker (Gibco); lane 4: V_H mAb 13AD; lane 5: V_L mAb 13AD; lane 6: assembled scFv 13AD. A 1.5% agarose gel is shown.

higher risk of cutting in antibody genes and they also incorporate internal restriction sites in the variable region genes that create mismatches with the antibody template and bias amplification by poor primer hybridization. Furthermore, AvrII, SacI and SpeI sites, which are present in most sets of enzymes used to date, cut in the majority of mouse λ chains and are therefore not suitable for simultaneous cloning of λ and κ light chains.

(vii) A frequently observed phenomenon is the contamination of antibody libraries with uncut recipient vectors (Courtney et al., 1995; Johansen et al., 1995). Normally antibody-free vectors have a growth advantage over scFv-encoding ones and cause problems during the enrichment of antigen-binding antibody sequences by phage display. Therefore, the pAK100 vector (Fig. 1E), with a tetracycline resistance cassette (tetA and tetR; 2101 kb) inserted between the two different SfiI sites is used as the recipient. Digestion with SfiI yields a linearized

vector which can be easily separated from the uncut one by gel electrophoresis. The loss of tet resistance can further ensure complete cutting of the recipient vector.

(viii) In order to avoid immunity to superinfection (Stengele et al., 1990), which is caused by expression of fusion proteins containing full length gIIIp, it is beneficial to use a truncated version of gIIIp. The truncated gIII 250–406 (Fig. 3B; Lowman et al., 1991), which is shorter than the more commonly used gIII 198–406 version, was chosen in order to eliminate a long glycine/serine rich linker stretch that favored instability. More importantly, the unpaired cysteine 201 at the end of the N-terminal domains, which reduces the folding yield of antibody-gIII fusions (data not shown) was also removed by this approach. Given that even background expression of truncated pIII fusions has been shown to be able to suppress superinfection to some extent (Ørum et al., 1993), it is furthermore an important

Table 2
Summary of cloned hybridomas

Hybridoma cell line	13AD	42PF	3D5	MOC31	3D7
Isotype	λ, IgG1	κ, IgG2b	κ, IgG2b	κ, IgG1	κ, IgG1
Tumor cell fusion partner	X63Ag8.653	X63Ag8.653	X63Ag8.653	X63Ag8.653	X63Ag8.653
Antigen	LZ	LZ(7P14P)	(his)₅ tag	EGP-2	DDT
Binders without panning	0/20	2/14	4/12	0/22	3/10
Binders after two panning rounds	5/6	nd	nd	8/10	nd
Identified aberrant or	aVH13AD.1	aVL42PF.λ	nd	nd	nd
non-binding sequences	aVH13AD.2				
PCR amplification by: Pharmacia Primer Mix					
V_L	No	No	nd	nd	No
V_H	Yes	Yes	nd	nd	Yes
Primers derived from Orlandi et al., 1989					
V_L	nd	nd	nd	No	nd
V_H	nd	nd	nd	aVHref	nd

Five hybridomas of three different isotypes have been cloned according to the scheme outlined in Fig. 1 and Fig. 9. For all hybridomas X63Ag8.653 (Kearney et al., 1979) was used as the tumor cell fusion partner. Hybridoma 13AD and 42PF produce antibodies directed against leucine zippers (Leder et al., 1995), 3D5 against C-terminal his tags (Lindner et al., 1997), MOC31 against the epithelial glycoprotein-2 (Souhami et al., 1988) and 3D7 against a derivative of DDT (Bürgisser et al., 1990). Functional binders (signal > 10 times background) are identified by phage ELISA as described in Section 2. The amino acid sequences of all identified aberrant chains are listed in Fig. 5. The sequences of aVH13AD.1 and aVHref are identical (except for amino acid 56; Fig. 5) to the aberrant V_H sequence published by Kütemeier et al. (1992). The same sequence was exclusively found (three independently sequenced clones) during V_H amplification of hybridoma MOC31 using primers derived from Orlandi et al. (1989), whereas V_L could not be amplified using this primer set. Amplification of V_L using the commercially available primer mix of Pharmacia failed in the case of hybridomas 42PF and 3D7, probably because appropriate sequences are absent from this mix, as well as for hybridoma 13AD, due to its λ isotype (nd = not determined; no = no PCR product detected; yes = PCR product detected).

A

pAK100scFv, pAK200scFv, pAK300scFv

```
                                                   end lacI
...ATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGC
...M  Q  L  A  R  Q  V  S  R  L  E  S  G  Q  *
```

```
                      tHP terminator
GGTACCCGATAAAAGCGGCTTCCTGACAGGAGGCCGTTTTGTTTTGCAGC
```

```
               CAP binding site
CCACCTCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGG
```

```
   -35          .          -10          lac-operator
CTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGA
                                         I-> mRNA
```

```
          SD1       LacZ
TAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTTCTAGA
                      M  T  M  I  T  N  F  *
```

```
   SD2          pelB signal sequence
TAACGAGGGCAAATCATGAAATACCTATTGCCTACGGCAGCCGCTGGATT
            M  K  Y  L  L  P  T  A  A  A  G  L
```

```
           SfiI              FLAG        VL
GTTATTACTCGCGGCCCAGCCGGCCATGGCGGACTACAAAGAY...
L  L  L  A  A  Q  P  A  M  A  D  Y  K  D  ...
```

pAK400scFv

```
   SD2                pelB signal sequence
...GAAGGAGATATACATATGAAATACCTATTGCCTACGGCAGCC...
   T7g10              M  K  Y  L  L  P  T  A  A  ...
```

B

pAK100scFv

```
   VH     SfiI     EcoRI         myc tag
...CGGCCTCGGGGGCCGAATTCGAGCAGAAGCTGATCTCTGAGGAAGAC
...  A  S  G  A  E  F  E  Q  K  L  I  S  E  E  D
```

```
          geneIII 250-406
CTGTAGGGTGGTGGCTCTGGTTCCGGTGATTTTGATTATGAAAAG...
L  *  G  G  G  S  G  S  G  D  F  D  Y  E  K  ...
```

pAK200scFv

```
   VH     SfiI        geneIII 250-406
...CGGCCTCGGGGGCCGAGGGCGGCGGTTCTGGTTCCGGTGATTTT...
...  A  S  G  A  E  G  G  G  S  G  S  G  D  Y  ...
```

pAK300scFv, pAK400scFv

```
   VH     SfiI      his tag
...CGGCCTCGGGGGCCGATCACCATCATCACCATCATTAGT...
...  A  S  G  A  D  H  H  H  H  H  H  *
```

C

pAK500scFv

```
   VH    SfiI    EcoRI              dHLX
.....CGGCCTCGGGGGCCGAATTCCCCAAACCTAGCACCCCCCCTGGCA
.....  A  S  G  A  E  F  P  K  P  S  T  P  P  G  S
```

```
GCAGTGGTGAACTGGAAGAGCTGCTTAAGCATCTTAAAGAACTTCTGAAG
S  G  E  L  E  E  L  L  K  H  L  K  E  L  L  K
```

```
GGCCCCCGCAAAGGCGAACTCGAGGAACTGCTGAAACATCTGAAGGAGCT
G  P  R  K  G  E  L  E  E  L  L  K  H  L  K  E  L
```

```
                          his tag
GCTTAAAGGTGGGAGCGGAGGCGCGCCGCACCATCATCACCATTGACGTC
L  K  G  G  S  G  G  A  P  H  H  H  H  H  *
```

```
HindIII
TAAGCTT...
```

pAK600scFv

```
   VH    SfiI    EcoRI    alkaline phosphatase (AP)
...CGGCCTCGGGGGCCGAATTCCGGACACCAGAAATGCCTGTTCTG...
   A  S  G  A  E  F  R  T  P  E  M  P  V  L  ...
                             start AP
```

```
                                      HindIII
...CTCTTCTACACCATGAAAGCCGCTCTGGGGCTGAAATAAGCTT...
...L  F  Y  T  M  K  A  A  L  G  L  K  *
                                    end AP
```

Fig. 3. *A*: upstream sequence of pAK100scFv, pAK200scFv, pAK300scFv and pAK400scFv. The region from the end of the *lacI* repressor gene to the beginning of the antibody V_L domain is shown. The *lacI* repressor gene, t_{HP} terminator sequence, CAP binding site, *lac* promotor/operator region (*lac p / o*) including the −35 and −10 sequence, Shine-Dalgarno (SD) sequence of *lacZ* (SD1), lacZ peptide, a second SD sequence (SD2), *pelB* signal sequence, N-terminal SfiI site, four amino acid FLAG tag and the start of the V_L domain (bold) are indicated above the sequence. For pAK400, the 15 bp upstream from the pelB start codon are replaced by a sequence including the SD sequence of the phage T7 gene10. *B*: downstream sequence of pAK100scFv, pAK200scFv, pAK300scFv and pAK400scFv. Relevant differences in the downstream sequences of pAK100, pAK200, pAK300 and pAK400 are shown. The last two bases of V_H (bold), SfiI and EcoRI restriction sites, myc or his tags and the start of geneIII(250-406) are indicated above the sequence. The stop codons are represented by asterisks. This corresponds to the region of the right-hand SfiI site in Fig. 4. *C*: sequences of EcoRI/HindIII fusion cassettes as used in pAK500 and pAK600. The dHLX dimerization motif was taken from Pack et al. (1993). The complete sequence of the mature *E. coli* alkaline phosphatase (AP) gene can be found in Shuttleworth et al. (1986). In order to provide a EcoRI/HindIII cloning cassette, the two internal EcoRI sites of the AP-gene have been removed by silent mutations (A. Knappik, unpublished data).

improvement to engineer the system such that complete product repression prior to helper phage infection can be ensured (see below).

(ix) A strong upstream t_{HP} terminator (Nohno et al., 1986) was incorporated between the *lacI* gene and the *lac* promoter region of pAK100 (Fig. 1E, Fig. 3A). This t_{HP} terminator sequence, in combina-

tion with glucose repression of the *lac* promoter (De Bellis and Schwartz, 1990), completely abolishes background expression before induction (for details see Krebber et al., 1996a). By these measures, selection against toxic scFv-gIII fusion proteins is avoided during propagation steps and plasmid maintenance is thus significantly improved.

(x) The *lac* repressor is encoded on the phagemid to ensure strain independent *lac* promoter repression.

(xi) Combining a synthetic SD sequence with a *pelB* signal sequence (Fig. 3A) leads to an only moderate level of translation allowing a low level of scFv-gIII expression upon induction by IPTG.
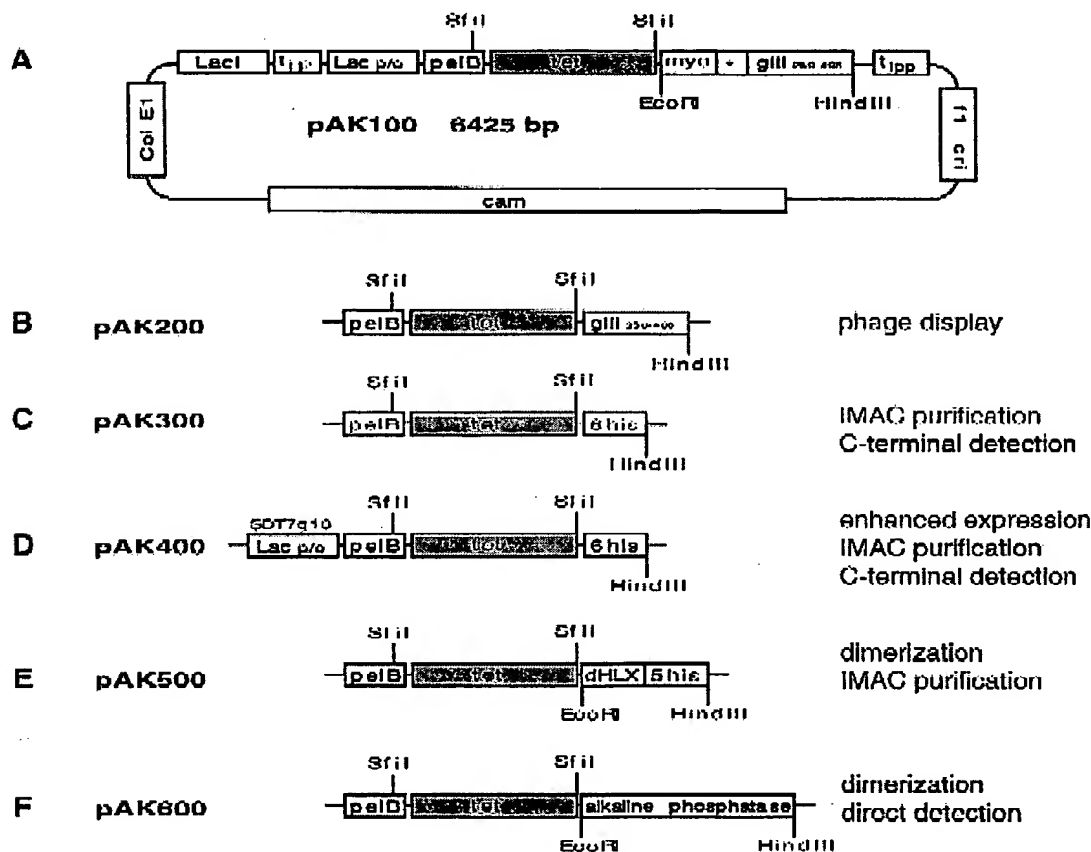
Fig. 4. pAK vector series. Phage display vector pAK100 (*A*) and related vectors (*B–F*) are useful to build modifications into antibody fragments cloned by the strategy outlined in Fig. 1 and Fig. 9. pAK200-pAK600 contain the same elements as described for pAK100 except for the modified cassettes shown. All vectors contain a *tet* resistance cassette (*tetA* and *tetR*; 2101 bp) to facilitate the monitoring of SfiI cutting. The in-frame fusion to geneIII250–406 using pAK100 (*A*) leads firstly into a myc-tag (Munro and Pelham, 1986) to be used as a detection handle, followed by an amber codon (asterisk). Depending on the strain used it is possible to switch between soluble expression of scFvs (in the case of non-suppressor strains such as JM83) and expression of scFv gene3 fusions (with suppressor strains such as XL1-Blue). The direct in-frame fusion to geneIII250–406 as in pAK200 (*B*) lacks the EcoRI site, myc-tag and amber codon. C-terminal his tag fusions can be used for purification by IMAC as well as for detection by an anti his tag antibody (Lindner et al., 1992, 1997; Kalinke et al., 1996) (*C*, *D*). Fusion partners, including helices for dimerization (Pack et al., 1993) (*E*) and alkaline phosphatase (AP) for direct detection of antigens by dimerized APscFv fusions (Lindner et al., 1997) (*F*), can be added to pAK100 or pAK400. The expression strength of either antibody or antibody fusions can be enhanced by replacing the original Shine Dalgarno (SD2) sequence by the stronger SDT7g10, as carried out in pAK400 (*D*).

(xii) The insertion of an amber codon upstream of ΔgeneIII in pAK100, as described by Hoogenboom et al. (1991) and Lowman et al. (1991), allows switching between expression of membrane anchored scFv-geneIII250–406 fusion proteins and soluble scFvs simply by changing the expression host.

A: VL lambda

CDR/FR				
VL42PF.λ	D A V V T Q E S A L T T S P G E T V T L T C R S S T G A V T T S N Y A N W V Q E K P D H L F T G L I G G T N N R A P G V P A
VL13AD	D A V V T Q E S A L T T S P G E T V T L T C R S S I G A V T T S N D A N W V Q E K P D H L F T G L I G G T N N R A P G V P A

| VL42PF.λ | R F S G S L L G D K A A L T I T G A Q T E D E A I Y F C A L W F S N H W V F G G G T K L T V L G |
| VL13AD | R F S G S L I G D K A A L T I T G A Q N E D E A K Y F C A L W Y S N H W V F G G G T K L T V L G |

B: VL kappa

| VL42PF.κ | D I V L I Q S P L S L P V S L G D Q A S I A C R S S Q S L V Q S N G E T Y L H W Y L Q K P G Q S P E L L I Y K V S N R F S G |
| aVLref | D I V L T Q S P A S L A V S L G Q R A T I S Y R A S K S V S T S - G Y S Y M H W N Q Q K P G Q P P R L L I Y L V S N L E S G |

| VL42PF.κ | V P D R F S G S G S G T D F T L K I S R V E A E D L G V Y F C S Q S T H V F G G G T K L E I K R |
| aVLref | V P A R F S G S G S G T D F T L N I H P V E E E D A A T Y Y C Q H I R E L T R S E G G P S W K · |

C: VH

VH42PF	E V Q L Q E S G G G L A K P G G S L K L S C G A S G F T F R S Y A M S W V R Q T P E R R L E W V A T I N T G G S Y T F Y P
VH13AD	E V Q R V E S G G G L V K P G G S P K L S C A A S G F T F S S S A M S W V R Q T P E K R L E W V A T I T S G G R F T Y Y P
aVH13AD.1	E V Q L Q Q S G P G L V A P S Q S L S I T C T V S G F S L T S Y G V H W V R Q P P G K G L E W L V V I W S D - G S T T Y N
aVHref	Q V Q L Q Q S G P G L V A P S Q S L S I T C T V S G F S L T S Y G V H W V R Q P P G K G L E W L V V I W S D - G T T T Y N
aVH13AD.2	Q V Q L Q Q S G A E L V R S G A S V K L S C T A S G F N I K D Y Y I Y W V K Q R P K Q G L E W I G W I D P E N G D T E C A

VH42PF	D S V K G R F T I S R D N A K N T L Y L Q M I S L R S E D T A M Y Y C V G G D H G S S L L A Y W G Q G T L V T V S A
VH13AD	D S V K G R F T I S R D N A K N T L Y L Q M S S L R S E D T A M Y Y C A I L Y D V Y Y Q R L Y W G Q G T T L T V S S
aVH13AD.1	S A L K S R L S I S K D N S K S Q V F L K M N S L Q T D D T A M Y Y C A R E P P T T Y V C L L G P R D S G H V S
aVHref	S A L K S R L S I S K D N S K S Q V F L K M N S L Q T D D T A M Y Y C A R E P P T T Y V C L L G P R D H G H R L L
aVH13AD.2	P N F Q G K A T V T A D T S S N T A S L Q F S S L T C E D T A V D Y C D S L V I T T Y - G L L G S R N L S H R L L

It must also be taken into account, however, that amber suppression is not complete and for this reason, an analogous phage display vector was constructed (pAK200; Fig. 4B), lacking the stop codon of pAK100. This leads to a higher proportion of displayed scFv-geneIII fusion protein, as monitored by ELISA and Western blot (data not shown). Direct competition of the same scFvs cloned into pAK100 or pAK200, however, reproducibly results in a complete enrichment of the pAK100 vector type after one round of phage panning (data not shown). This suggests that the lower level of fusion protein expression is a selective advantage.

(xiii) Chloramphenicol resistance (Kang et al., 1991) was used as the selective marker because it was found to allow more stringent selection than attainable with ampicillin, since the resistance protein does not leak into the medium. Furthermore, the use of chloramphenicol is advantageous over that of kanamycin or tetracycline because it does not reduce the phage titer as much (Johansen et al., 1995; Krebber et al., 1995; Krebber A., unpublished observations).

(xiv) The procedure can be used for library cloning, e.g. the repertoire from immunized mice, as well as for cloning of single sequences from hybridomas in a similar way. Furthermore, it is directly compatible with the recently introduced selectively infective phage system (SIP), which allows in vivo and in vitro selection of cognate protein/ligand interactions by strictly coupling the infectivity of filamentous phages to the binding event (Krebber et al., 1995; Krebber et al., 1996b).

(xv) Since the optimized phage display vector pAK100 is engineered to achieve low levels of expression, it is not a useful large scale production system for well folding and soluble antibody fragments. Thus, a compatible high-level expression plasmid has also been engineered. (Fig. 3A, Fig. 4D).

(xvi) A compatible vector series that facilitates various modifications of scFv fragments subsequent to cloning into pAK100 is also available (Fig. 3B, Fig. 3C, Fig. 4C–F; see also Ge et al., 1995).

3.2. Amplification of V region genes and assembly into the scFv format

The V_L back primer mix (LB1-17 and LBλ, representing a total of 131 variants) paired with five V_L forward primers (LF1, 2, 4, 5 and LFλ) and the V_H back mix (HB1-19, representing a total of 94 variants) paired with four V_H forward primers (HF1-4) have been used to amplify V_L and V_H domains from a variety of antibody cDNAs (Table 2).

Our improved primer set (Table 1) has been tested in different laboratories on cDNA derived from 12 hybridoma cell lines of different specificities and family sub-types to date. In all cases, the first PCR amplification yielded sufficient amounts of products for cloning, with a sharp band at the predicted size of 375–402 bp for V_L or 386–440 bp for V_H. Typical examples of V_H and V_L genes amplified from cDNA of two hybridomas, 42PF and 13AD, which secrete monoclonal antibodies directed against leucine zippers (Leder et al., 1995), are shown (Fig. 2). Using the same cDNA in combination with a commercially available primer mix (recombinant phage antibody system (Pharmacia)) or primers derived from Orlandi et al. (1989), amplification of V_L failed in several cases (Table 2), underlining the importance of an extended primer mix.

Fig. 5. Sequence alignment of functional and aberrant variable domains expressed by the hybridoma cell lines 13AD and 42PF. Residue numbers are according to Kabat et al. (1991). The 7 amino acids at each end are encoded by the PCR primer sequences. *A*: V_λ amino acid sequences. VL42PF.λ: non-binding V_λ found in clone 42PF; identical to germline $V_\lambda 1$ sequence (Weiss and Wu, 1987) except for F92Y. VL13AD (X99507): functional, antigen-binding V_λ sequence of hybridoma 13AD. *B*: V_κ amino acid sequences. VL42PF.κ (X99509): functional, antigen-binding sequence of clone 42PF. aVLref: aberrant V_κ transcript found in P3X63Ag8.653 (Carroll et al., 1988 (M35669); Duan and Pomerantz, 1994; Cabilly and Riggs, 1985; Strohal et al., 1987; Yamanaka et al., 1995). *C*: V_H amino acid sequences. VH42PF (X99508): functional V_H of hybridoma 42PF. VH13AD (X99506): functional V_H of hybridoma 13AD. aVH13AD.1: aberrant $V_H.1$ found in clone 13AD, showing a frameshift in CDR3. aVHref: non-functional V_H published by Yamanaka et al. (1995); this sequence is identical to the unpublished result of Mocikat (D50398). aVH13AD.2: aberrant $V_H.2$ found in 13AD. Sequence is different to aVH13AD.1, and also contains a frameshift in CDR3. EMBL accession numbers are given in brackets.

For further analysis the V_L and V_H PCR products of 42PF and 13AD have been cloned separately into the pCR-Script vector (Stratagene) and sequenced. For 42PF two plausible light chain sequences devoid of frameshifts, stop codons, deletions or atypical amino acids for murine V_L domains (VL42PF.κ, VL42PF.λ) were found, together with one heavy chain sequence (VH42PF) (Fig. 5). For hybridoma 13AD only 3 of 57 clones analyzed contained a bona fide functional heavy chain gene, denoted VH13AD, whereas two additional non-functional heavy chain sequences aVH13AD.1 (five clones) and aVH13AD.2 (49 clones) were found (Fig. 5). Both heavy chains are aberrantly rearranged at the DJ recombination site in CDR3 and contain several framework amino acids which deviate from the observed consensus of antibody sequences (Fig. 5). Sequencing of five V_L chains, amplified exclusively by the λ primer pair LBλ/LFλ, yielded a unique sequence denoted VL13AD (Fig. 5). Thus, both the 13AD and 42PF hybridoma produced more than one PCR-amplifiable heavy or light chain.

As outlined in Fig. 1, all amplified V_L and V_H domains have been linked by SOE-PCR, as shown for 13AD and 42PF (Fig. 2). These were subsequently digested by SfiI and ligated into the improved phage display vector pAK100.

3.3. Screening and enrichment of functional scFv sequences derived from hybridomas

After transformation of the ligation reaction into the recombination deficient *E. coli* strain XL1-Blue, 10–22 individual colonies were grown separately and infected by helper phage as described in Section 2. The recombinant scFvs, displayed on the surface of filamentous phage, were tested for antigen binding in a typical phage ELISA. In those cases where the parental hybridoma cell line did not produce large amounts of contaminating, non-functional light or heavy chain, about one third of the screened colonies contained the sequence information of the binding scFv fragments (examples are hybridoma 42PF amplified with a V_L κ mix, devoid of λ primers and hybridomas 3D5 and 3D7; Table 2). At the other extreme, an initial screening of phages derived from individual colonies of 13AD did not yield any functional binders. As previously demonstrated by the

sequencing of individual V_H domains, functional sequences are greatly diluted by aberrant heavy chains in this hybridoma cell line. In order to identify and enrich functional binders, 10^5 *E. coli* colonies were pooled after transformation and subjected to two rounds of phage panning. After each round, six clones were tested for antigen binding in a phage ELISA. Two of six and five of six clones from the first and second panning rounds, respectively, were found to be positive for antigen binding. All positive clones had identical sequences (VL13AD paired with VH13AD), whereas all non-binding scFv sequences contained the aberrant heavy chains aVH13AD.1 or aVH13AD.2, occasionally in combination with point mutations in the light chain gene.

Clones which were found to be positive in phage
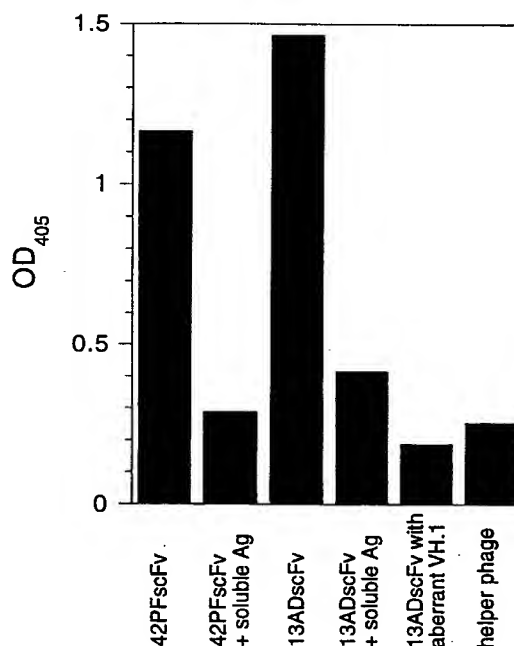
Fig. 6. Competition phage ELISA (13AD, 42PF). Competition phage ELISA with phages displaying functional and non-functional scFv fragments derived from hybridomas 13AD and 42PF. The ELISA was performed as described in Section 2. The non-binding 13ADscFv clone contains the aberrant aVH13AD.1 chain (Fig. 5C) and the functional VL13AD chain (Fig. 5A). For inhibition, phages were preincubated for 10 min with 10^{-4} M soluble peptide antigen before applying the mixture to the antigen-coated plate. As a negative control, an assay with VCS helper phage was performed.
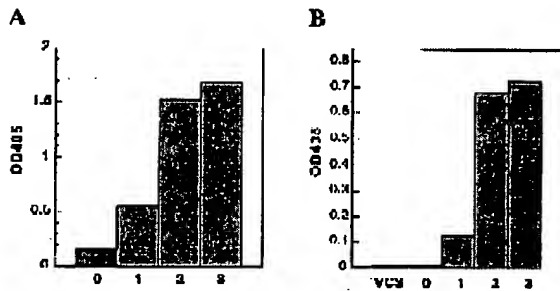
Fig. 7. Enrichment by panning. *A*: cloning of hybridoma MOC31. Enrichment of EGP-2 binding scFv by phage panning. *B*: repertoire cloning. Enrichment of ampicillin binding scFv displaying phages derived from anti-ampicillin library I. Phage pools $(5 \cdot 10^{10}$ cfu/well) prepared after 0, 1, 2 and 3 rounds of phage panning (column label), as well as VCS helper phage as a negative control, were tested for specific antigen binding in a phage ELISA as described in Section 2. Selective enrichment was also indicated by an increased number of phages eluted in subsequent panning rounds (data not shown).

ELISAs were further characterized by antigen inhibition studies to verify that the binding was antigen-specific (Fig. 6). In the case of hybridoma MOC31 (Souhami et al., 1988), again no binders were initially obtained, and two to three rounds of phage panning were required to enrich binding scFv fragments to a level that allowed identification of functional antibody sequences in individual clones (Fig. 7A; Table 2). This shows that the relevant sequences of numerous 'monoclonal' antibodies can be hidden in a pool of closely related antibody-like sequences and that, in the absence of panning, rigorous testing would be required in order to identify the correct sequence.

3.4. Cloning of the antibody response from immunized mice

The procedure described for hybridoma cloning was also applied to mRNA isolated from spleen cells of an immunized mouse. In addition, B-cells from the same mouse were fused to the tumor cell line X63Ag8.653 as in the case of monoclonal antibody production, but were kept as a pool for 10 days. This pool of hybridomas was subsequently used as a source of mRNA. The latter experiment was carried out because it might seem conceivable that B cells which have been stimulated by the antigen fuse

preferentially (Köhler and Milstein, 1976). From fusion experiments, only a small number of a few thousand candidate clones is typically obtained, of which a high proportion usually codes for antigen binding antibody sequences. Since productive pairs of V_L and V_H domains are separated during the cloning process and are subsequently combined randomly to form scFvs, fairly large libraries are necessary to ensure that all original V_L and V_H pairings are represented (Gherardi and Milstein, 1992; Posner et al., 1994). A comparison of anti-ampicillin libraries derived from fused (library I) and unfused B-cells (library II) of the same immunized mouse should determine whether cell fusion prior to mRNA preparation is an advantageous enrichment step which enhances the probability of restoring functional V_L/V_H pairings in a small library. As outlined in Table 3, both libraries contained binding scFv fragments which could be enriched with a similar efficiency after two or three rounds of panning (Fig. 7B). Sequencing revealed that the same sequences were isolated simultaneously from both libraries (data not shown), indicating that B cell fusion to tumor

Table 3

Selection of ampicillin binding scFv fragments from B cell repertoires

	Library I	Library II
Source of RNA	Spleen cells fused to tumor cell line X63Ag8.653	B cells
Antigen	Ampicillin	Ampicillin
Library size	$4 \cdot 10^6$	$6 \cdot 10^6$
Clones containing SfiI insert	20/20	20/20
Clones expressing scFv	22/30	18/30
Binders before panning	0/12	0/12
Binders after two rounds	5/12	nd
Binders after three rounds	19/24	9/12

Spleen cells derived from the same BALB/c mouse immunized with ampicillin were taken for library construction before (library II) and after (library I) fusion to the tumor cell line X63Ag8.653. Both libraries were transformed into XL1-Blue cells by electroporation. The amount of SfiI insert-containing clones in the initial library was monitored at the DNA level by restriction analysis whereas the amount of full length scFv-expressing clones was analyzed by Western blot analysis, using the N-terminal FLAG detection system combined with C-terminal myc tag detection (data not shown). Binding scFv fragments were identified by phage ELISA. The enrichment process was followed by ELISA using phage pools as shown in Fig. 7B.

cells prior to mRNA preparation, at least in our experience, has no significant beneficial influence on library composition.

3.5. Soluble expression and modification of cloned scFv fragments

A scFv fragment obtained from the anti-ampicillin library I (scFvaL2) was sub-cloned into pAK300 and pAK400 (Fig. 4) for soluble expression in JM83. In comparison with the low expression medium (LE medium) used for phage display, changing to an expression medium devoid of glucose immediately increases the expression level of recombinant protein without any modifications to the vector system (De Bellis and Schwartz, 1990). Changing the translation initiation region present in pAK100 or pAK300 into a much stronger Shine-Dalgarno sequence (SDT7g10), such as that present in pAK400 (Fig. 3A, Plückthun et al., 1996), results in a further significant enhancement of protein expression. As shown in Fig. 8, the expression level strongly influences the ratio of soluble to insoluble scFv protein. While at the lower expression level of pAK300scFvaL2 100% of the scFv is soluble and

c   pAK300scFvaL2

c   pAK400scFvaL2

    marker

I   pAK300scFvaL2

s   pAK300scFvaL2

I   pAK400scFvaL2
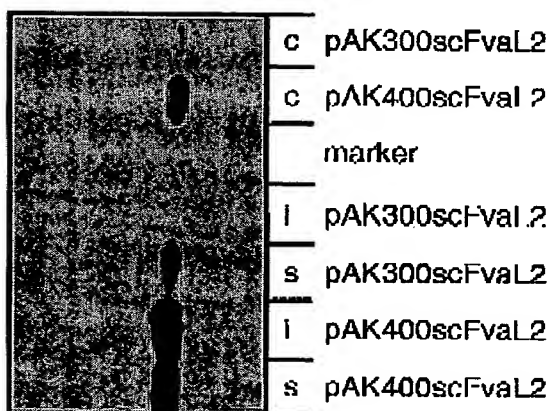
s   pAK400scFvaL2

Fig. 8. Enhanced expression of aL2 in pAK400. The scFvaL2 was expressed in JM83 harboring pAK300scFvaL2 or pAK400scFvaL2 (Fig. 3 and Fig. 4). Expression levels were monitored by Western blot analysis as described in Section 2. c: whole culture (soluble fraction, insoluble fraction and culture supernatant), where the loaded sample corresponds to 1 ml of culture OD_{600} of 0.01; i: insoluble fraction; and s: soluble fraction, where the loaded sample corresponds to 1 ml culture at an OD_{600} of 0.05.
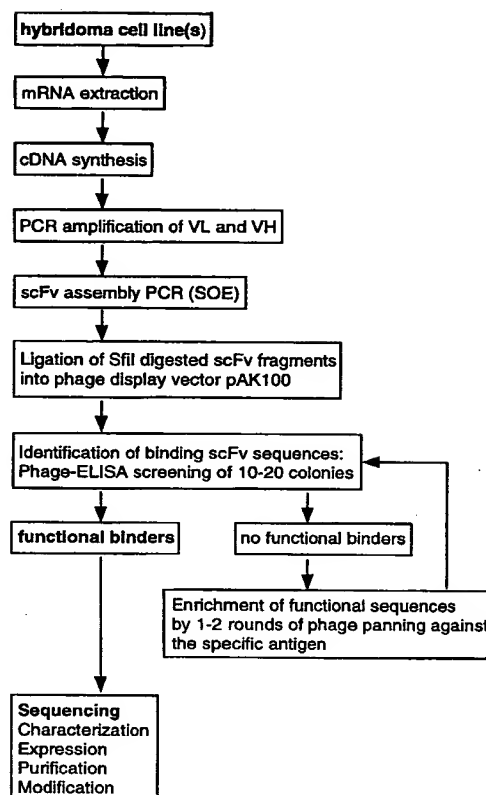
Fig. 9. Outline: Generation of scFv antibodies from hybridomas. A flow diagram summarizing the most important steps is shown.

functional, the enhanced expression in pAK400scFvaL2 causes production of more soluble but also large amounts of insoluble material (Fig. 8).

The scFvaL2 has comparatively favorable folding properties (A. Krebber, unpublished). For scFv fragments, which are already mainly found in the insoluble fraction after expression in pAK100 or pAK300, sub-cloning into pAK400 does not improve the yield of functional antibody fragment (data not shown). Instead, cell lysis problems, caused by such poorly folding proteins, are enhanced in the stronger expression vector since higher expression levels normally lead to a higher proportion of non-functional aggregates (Le Calvez et al., 1995), which are likely to impair growth of the expression host. Therefore, it has proved to be advantageous to adapt the expression level to the particular scFv sequence which has

to be expressed by the choice of vector and induction conditions.

Moreover, SfiI cassettes of scFv fragments can be fused directly in frame with oligo-histidine tags for purification by IMAC (Lindner et al., 1992), with dimerization or tetramerization modules to obtain dimeric or multimeric scFv antibodies (Pack et al., 1993, 1995) and with enzymes such as alkaline phosphatase to produce dimeric scFv molecules which can be detected directly by virtue of their enzymatic activity (Lindner et al., 1997; Fig. 3B and C and Fig. 4). If IMAC does not directly yield pure scFv antibody, an immunoaffinity column using an anti-FLAG mAb can also be employed (Kalinke et al., 1996). Combination of the N-terminal FLAG tag and the C-terminal myc or his tags allows monitoring of full length scFv product formation, since antibodies against all three tags are available (Munro and Pelham, 1986; Knappik and Plückthun, 1994; Lindner et al., 1997). Therefore, Western blot detection of N- and C-terminal degradation and of proteolysis of particular scFv sequences becomes easily possible. The whole spectrum of compatible modification cassettes (see also Ge et al., 1995; Plückthun et al., 1996) combined with the pAK vector series creates a highly versatile system, allowing easy characterization and further genetic engineering of scFv fragments initially obtained.

4. Discussion

The improved phage display system based on the pAK vector series (Fig. 4), an extended primer mix (Table 1) and a very straightforward cloning procedure (Fig. 1) proved to be robust and reliable both in a library setting and for hybridoma cloning. Following the scheme outlined in Fig. 9 all hybridomas tested to date could be cloned, characterized for functional antigen binding and sequenced with a reasonable effort, in as few as 10 days (hybridoma 3D5).

The optimized phage display system was suitable for eliminating high amounts of non-functional chains that are transcribed from various aberrant mRNAs in some hybridoma cell lines. In contrast to other methods (Nicholls et al., 1993; Duan and Pomerantz, 1994; Ostermeier and Michel, 1996), only functional

and binding antibody genes will be sequenced. After RNAseH treatment of aberrant RNA/DNA hybrids (Ostermeier and Michel, 1996), eight out of 12 sequenced clones were still derived from aberrant pseudogenes, whereas without mRNA treatment all nine clones tested carried pseudogenes. Duan and Pomerantz (1994) used ribozyme treatment to improve the ratio of aberrant to functional sequences from three positive clones in 150 to 12–34 in 150. Both methods depend on the availability of sequence information of the aberrant chain prior to cloning, whereas phage display simply enriches binding sequences over any kind of contaminating chain. This has been found to be particularly important, as both of the hybridoma cell lines we have characterized in detail (13AD and 42PF) produced aberrant mRNAs which seem to be specific for the individual hybridomas and could not be found in the published literature or any database (Fig. 5). The origin of some of the aberrant mRNAs could not be traced to the myeloma cell lines originally utilized for cell fusion. Hybridoma 13AD, for example, contained three different heavy chain mRNAs, of which only one is known to be derived from the tumor cell line X63Ag8.653. It seems plausible that many additional non-binding chains originate from aberrant rearrangements of the second allele of the B-cell involved in the generation of the particular hybridoma. Termination of rearrangement in the immunoglobulin loci takes place only after synthesis of a functional membrane-bound immunoglobulin. Thus, a given B-cell may produce aberrant mRNAs that contain stop codons or frameshifts in addition to the functional mRNA that is translated into the mature immunoglobulin chain. Furthermore, some hybridomas may be the result of the fusion of more than one B-cell with the myeloma cell. This gives rise to the possibility that more than one typical heavy and light chain gene is expressed, as observed for hybridoma 42PF. Such typical in-frame but non-binding sequences cannot be distinguished from the binding chain by sequencing. This underlines the importance of a functional test involving antigen binding in order to avoid the risk of isolating an incorrect sequence.

When this is taken in combination with possible PCR errors or mutations introduced by the primer mix, it can be envisaged how cloning of a hybridoma

can very easily generate a diverse collection of different scFv fragments. Thus, functional screening is always superior to sequence analysis of individual clones.

If more than one monoclonal antibody against the same antigen is available, pooled hybridoma cell lines can be used as a source for mRNA extraction, as carried out in the case of the anti-ampicillin library I. This allows for fast and inexpensive simultaneous cloning of several different antibodies in one experimental setup. Alternatively, a small fraction of B-cells prepared for traditional monoclonal antibody generation can be set aside for phage library construction. It has been reported, however, that parallel screening of hybridomas and phage libraries, as described by Gherardi and Milstein (1992); Kettleborough et al. (1994) and Ames et al. (1995), has led to the discovery of different antibody sequences from the two sources. This may simply reflect the fact that neither hybridoma generation nor phage libraries provides an exhaustive sampling of the immune response. The findings may suggest in addition that recombinant expression of antibody fragments combined with phage display selects against certain antibody sequences (see also Riechmann and Weill, 1993; Posner et al., 1994; Jackson et al., 1995). The absence of certain sequences in phage libraries might be due to insufficient library size, PCR amplification of only a subset of binding variable genes or selection against phagemids expressing less stable or less well folding antibody fragments, particularly if they are incorporated into phage particles less frequently or impair growth of *E. coli* (Knappik and Plückthun, 1995; Krebber et al., 1996a). We believe, however, that our optimized cloning procedure and the tightly regulatable phage display vector will contribute to overcoming some of these biases and will therefore facilitate the construction of more diverse antibody libraries. Given the stress that antibodies fused with gIIIp impose on the cell, a high expression level is of no importance and actually serves as a burden for phage display. In contrast, the absence of background expression before induction is of utmost importance if loss of clones from the library or gene deletions are not to quickly accumulate. Thus, the phage display vector pAK100 was optimized by lowering the expression level for phage display and, more importantly, by eliminating background expres-

sion before helper phage infection (Krebber et al., 1996a). In contrast to many previous vector constructs, therefore, even in a library setting, no deletions, empty vectors, recombination events or other symptoms of instability have been detected to date. Moreover we claim that it is advisable to sub-clone selected scFvs into related, but more powerful, expression vectors subsequent to cloning, since unbiased library screening by phage display and maximized functional production of a particular antibody fragment are likely to require different expression optima.

Acknowledgements

References

Ames, R.S., Tornetta, M.A., McMillan, L.J., Kaiser, K.F., Holmes, S.D., Appelbaum, E., Cusimano, D.M., Theisen, T.W., Gross, M.S., Jones, C.S., Silverman, C., Porter, T.G., Cook, R.M., Bennett, D. and Chaiken, I.M. (1995) Neutralizing murine monoclonal antibodies to human IL-5 isolated from hybridomas and a filamentous phage Fab display library. J. Immunol. 154, 6355.

Barbas III, C.F. and Wagner, J. (1995) Synthetic human antibodies: Selecting and evolving functional proteins. Methods Companion Methods Enzymol. 8, 94.

Barbas III, C.F., Kang, A.S., Lerner, R.A. and Benkovic, S.J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: The geneIII site. Proc. Natl. Acad. Sci. USA 88, 7978.

Berger, S.L. and Chirgwin, J.M. (1989) Isolation of RNA. Methods Enzymol. 180, 3.

Bürgisser, D., Frey, S., Gutte, B. and Klauser, S. (1990) Preparation and characterization of polyclonal and monoclonal antibodies against the insecticide DDT. Biochem. Biophys. Res. Commun. 166, 1228.

Cabilly, S. and Riggs, A.D. (1985) Immunoglobulin transcripts and molecular history of a hybridoma that produces antibody to carcinoembryonic antigen. Gene 40, 157.

Carroll, W.L., Mendel, E. and Levy, S. (1988) Hybridoma fusion cell lines contain an aberrant kappa transcript. Mol. Immunol. 25, 991.

Clackson, T., Hoogenboom, H.R., Griffiths, A.D. and Winter, G.

(1991) Making antibody fragments using phage display libraries. Nature 352, 624.

Courtney, B.C., Williams, K.C. and Schlager, J.J. (1995) A phage display vector with improved stability, applicability and ease of manipulation. Gene 165, 139.

De Bellis, D. and Schwartz, I. (1990) Regulated expression of foreign genes fused to *lac*: control by glucose levels in growth medium. Nucleic Acids Res. 18, 1311.

Deng, S., MacKenzie, C.R., Sadowska, J., Michniewicz, J., Young, N.M., Bundle, D.R. and Narang, S.A. (1994) Selection of antibody single-chain variable fragments with improved carbohydrate binding by phage display. J. Biol. Chem. 269, 9533.

Desplancq, D., King, D.J., Lawson, A.D.G. and Mountain, A. (1994) Multimerization behaviour of single chain Fv variants for the tumor-binding antibody B72.3. Prot. Eng. 7, 1027.

Dower, J.W., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. 16, 6127.

Duan, L. and Pomerantz, R.J. (1994) Elimination of endogenous aberrant kappa chain transcripts from sp2/0-derived hybridoma cells by specific ribozyme cleavage: utility in genetic therapy of HIV-1 infections. Nucleic Acids Res. 22, 5433.

Dziegiel, M., Nielson, L.K., Anderson, P.S., Blancher, A., Dickmeiss, E. and Engberg, J. (1995) Phage display used for gene cloning of human recombinant antibody against the erythrocyte surface antigen, rhesus D. J. Immunol. Methods 182, 7.

Freund, C., Ross, A., Plückthun, A. and Holak, T.A. (1994) Structural and dynamic properties of the Fv fragment and the single-chain Fv fragment of an antibody in solution investigated by heteronuclear three-dimensional NMR spectroscopy. Biochemistry 33, 3296.

Ge, L., Knappik, A., Pack, P., Freund, C. and Plückthun, A. (1995) Expressing antibodies in *Escherichia coli*. In: C.A.K. Borrebaeck (Ed.), Antibody Engineering, 2nd edn. Oxford University Press, Oxford, p. 229.

Gherardi, E. and Milstein, C. (1992) Original and artificial antibodies. Nature 357, 201.

Glockshuber, R., Malia, M., Pfitzinger, I. and Plückthun, A. (1990) A comparison of strategies to stabilize immunoglobulin Fv-fragments. Biochemistry 29, 1362.

Güssow, D. and Seemann, G. (1991) Humanization of monoclonal antibodies. Methods Enzymol. 203, 99.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D., Hudson, P. and Winter, G. (1991) Multi-subunit proteins on the surface of filamentous phage: Methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. 19, 4133.

Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. and Pease, L.R. (1989) Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. Gene 77, 61.

Huston, J.S., McCartney, J., Tai, M.S., Mottola-Hartshorn, C., Jin, D., Warren, F., Keck, P. and Oppermann, H. (1993) Medical applications of single-chain antibodies. Intern. Rev. Immunol. 10, 195.

Huston, J.S., George, A.J.T., Tai, M.-S., McCartney, J.E., Jin, D., Segal, D.M., Keck, P. and Oppermann, H. (1995) Single-chain Fv design and production by preparative folding. In: C.A.K.

Borrebaeck (Ed.), Antibody Engineering, 2nd edn. Oxford University Press, Oxford, p. 185.

Jackson, J.R., Sathe, G., Rosenberg, M. and Sweet, R. (1995) In vitro antibody maturation. J. Immunol. 154, 3310.

Johansen, L.K., Albrechtsen, B., Andersen, H.W. and Engberg, J. (1995) pFab60: a new, efficient vector for expression of antibody Fab fragments displayed on phage. Prot. Eng. 8, 1063.

Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M., Gottesman, K.S. and Foeller, C. (1991) Sequences of Proteins of Immunological Interest, 5th edn. US Department of Health and Human Services, Public Service, NIH, Washington.

Kalinke, U., Krebber, A., Krebber, C., Bucher, E., Plückthun, A., Zinkernagel, R.M. and Hengartner, H. (1996) Monovalent single-chain Fv fragments and bivalent miniantibodies bound to vesicular stomatitis virus (VSV) protect against lethal infection. Eur. J. Immunol. 26, 2801.

Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J. and Lerner, R.A. (1991) Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc. Natl. Acad. Sci. USA 88, 4363.

Kaluza, B., Betzl, G., Shao, H., Diamantstein, T. and Weidle, U.H. (1992) A general method for chimerization of monoclonal antibodies by inverse polymerase chain reaction which conserves authentic N-terminal sequences. Gene 122, 321.

Kearney, J.F., Radbruch, A., Liesegang, B. and Rajewski, K. (1979) A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. J. Immunol. 123, 1548.

Kettleborough, C.A., Saldanha, J., Ansell, K.H. and Bendig, M.M. (1993) Optimization of primers for cloning libraries of mouse immunoglobulin genes using the polymerase chain reaction. Eur. J. Immunol. 23, 206.

Kettleborough, C.A., Ansell, K.H., Allen, R.W., Rosell-Vives, E., Güssow, D.H. and Bendig, M.M. (1994) Isolation of tumor cell-specific single-chain Fv from immunized mice using phage-antibody libraries and the re-construction of whole antibodies from these antibody fragments. Eur. J. Immunol. 24, 952.

Knappik, A. and Plückthun, A. (1994) An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. BioTechniques 17, 754.

Knappik, A. and Plückthun, A. (1995) Engineered turns of a recombinant antibody improve its in vivo folding. Prot. Eng. 8, 81.

Köhler, G. and Milstein, C. (1976) Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. Eur. J. Immunol. 6, 511.

Krebber, C., Spada, S., Desplancq, D. and Plückthun, A. (1995) Co-selection of cognate antibody-antigen pairs by selectively-infective phages. FEBS Lett. 377, 227.

Krebber, A., Burmester, J. and Plückthun, A. (1996a) Inclusion of an upstream transcriptional terminator in phage display vectors abolishes background expression of toxic fusions with coat protein g3p. Gene 178, 71.

Krebber, C., Spada, S., Desplancq, D., Krebber, A., Ge, L., Lindner, B. and Plückthun, A. (1996b) Selectively-infective

phage (SIP): a mechanistic dissection of a novel in vivo selection for protein ligand interactions. Submitted.

Kütemeier, G., Harloff, C. and Mocikat, R. (1992) Rapid isolation of immunoglobulin variable genes from cell lysates of rat hybridomas by polymerase chain reaction. Hybridoma 11, 23.

Kwak, J.-W., Choi, B.-K., Lee, D.-I., Kang, Y.-K., Seo, Y.-G., Cho, W.-K. and Han, M.H. (1996) Cloning and characterization of cDNAs coding for heavy and light chains of a monoclonal antibody (MabB23) specific for human plasma apolipoprotein B-100. Gene 169, 237.

Le Calvez, H., Fieschi, J., Green, J.M., Marchesi, N., Chauveau, J. and Baty, D. (1995) Paratope characterization by structural modelling of two anti-cortisol single-chain variable fragments produced in E. coli. Mol. Immunol. 32, 185.

Leder, L., Berger, C., Bornhauser, S., Wendt, H., Ackermann, F., Jelesarov, I. and Bosshard, H.R. (1995) Spectroscopic, calorimetric, and kinetic demonstration of conformational adaption in peptide-antibody recognition. Biochemistry 34, 16509.

Lindner, P., Guth, B., Wülfing, C., Krebber, C., Steipe, B., Müller, F. and Plückthun, A. (1992) Purification of native proteins from the cytoplasm and periplasm of Escherichia coli using IMAC and histidine tails: A comparison of proteins and protocols. Methods Companion Methods Enzymol. 4, 41.

Lindner, P., Bauer, K., Krebber, A., Nieba, L., Kremmer, E., Krebber, C., Honegger, A., Kinger, B., Mocikat, R. and Plückthun, A. (1997) Specific detection of his-tagged proteins with recombinant anti-his tag scFv-phosphatase or scFv-phage fusions. BioTechniques, in press.

Lowman, H.B., Bass, S.H., Simpson, N. and Wells, J.A. (1991) Selecting high-affinity binding proteins by monovalent phage display. Biochemistry 30, 10832.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D. and Winter, G. (1991) By-passing immunization. Human antibodies from V-gene libraries displayed on phage. J. Mol. Biol. 222, 581.

Marks, J.D., Griffiths, A.D., Malmqvist, M., Clackson, T.P., Bye, J.M. and Winter, G. (1992) By-passing immunization: Building high affinity human antibodies by chain shuffling. Biotechnology 10, 779.

McCafferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348, 552.

Miller, C.E., Huppi, K., Siwarski, D., Karpas, A., Newman, A., Mainhart, C. and Glaudemans, C.P.J. (1995) A murine antibody to Shigella dysenteriae type 1 employs V-genes that contain a rearranged codon for the λ light chain. Mol. Immunol. 32, 679.

Munro, S. and Pelham, H.R. (1986) An Hsp70-like protein in the ER: Identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell 46, 291.

Nicholls, P.J., Johnson, V.G., Blanford, M.D. and Andrew, S.M. (1993) An improved method for generating single-chain antibodies from hybridomas. J. Immunol. Methods 165, 81.

Nilsson, B. (1995) Antibody engineering. Curr. Opin. Struct. Biol. 5, 450.

Nohno, T., Saito, T. and Hong, J.-S. (1986) Cloning and complete nucleotide sequence of the Escherichia coli glutamine permease operon (glnHPQ). Mol. Gen. Genet. 205, 260.

Ohlin, M., Owman, H., Mach, M. and Borrebaeck, C.A.K. (1996) Light chain shuffling of a high affinity antibody results in a drift in epitope recognition. Mol. Immunol. 33, 47.

Orlandi, R., Güssow, D.H., Jones, P.T. and Winter, G. (1989) Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA 86, 3833.

Ørum, H., Andersen, P.S., Øster, A., Johansen, L.K., Riise, E., Bjørnvad, M., Svendsen, I. and Engberg, J. (1993) Efficient method for constructing comprehensive murine Fab antibody libraries displayed on phage. Nucleic Acids Res. 21, 4491.

Ostermeier, C. and Michel, H. (1996) Improved cloning of antibody variable regions from hybridomas by an antisense-directed RNAse H digestion of the P3-X63-Ag8.653 derived pseudogene mRNA. Nucleic Acids Res. 24, 1979.

Ostermeier, C., Essen L.-O. and Michel, H. (1995) Crystals of an antibody Fv fragment against an integral membrane protein diffracting to 1.28 Å resolution. Proteins 21, 74.

Pack, P., Kujau, M., Schroeckh, V., Knüpfer, U., Wenderoth, R., Riesenberg, D. and Plückthun, A. (1993) Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of Escherichia coli. Biotechnology 11, 1271.

Pack, P., Müller, K., Zahn, R. and Plückthun, A. (1995) Tetravalent miniantibodies with high avidity assembling in Escherichia coli. J. Mol. Biol. 246, 28.

Plückthun, A. (1994) Recombinant antibodies. In: C.J. van Oss and M.H.V. van Regenmortel (Eds.), Immunochemistry. Marcel Dekker, New York, p. 201.

Plückthun, A., Krebber, A., Krebber, C., Horn, U., Knüpfer, U., Wenderoth, R., Nieba, L., Proba, K. and Riesenberg, D. (1996) Producing antibodies in Eschericia coli: From PCR to fermentation. In: J. McCafferty and H.R. Hoogenboom (Eds.), Antibody Engineering: A Practical Approach. IRL Press, Oxford, p. 203.

Posner, B., Smiley, J., Lee, I. and Benkovic, S. (1994) Catalytic antibodies: Perusing combinatorial libraries. Trends Biochem. Sci. 19, 145.

Rees, A.R., Staunton, D., Webster, D.M., Searle, S.J., Henry, A.H. and Pederson (1994) Antibody design: beyond the natural limits. Trends Biotechnol. 12, 199.

Reiter, Y., Brinkmann, U., Kreitmann, R.J., Jung, S-H., Lee, B. and Pastan, I. (1994) Stabilization of the Fv fragments in recombinant immunotoxins by disulfide bonds engineered into conserved framework regions. Biochemistry 33, 5451.

Riechmann, L. and Weill, M. (1993) Phage display and selection of a site-directed randomized single-chain antibody Fv fragment for its affinity improvement. Biochemistry 32, 8848.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shuttleworth, H., Taylor, J. and Minton, N. (1986) Sequence of the gene for alkaline phosphatase from Escherichia coli JM83. Nucleic Acids Res. 14, 8689.

Souhami, R.L., Beverley, P.C.L. and Bobrow, L.G. (1988) Proceedings of the First International Workshop on Small-Cell Lung-Cancer Antigens. Lung Cancer 4, 1.

Stengele, I., Bross, P., Garcés, X., Giray, J. and Rasched, I. (1990) Dissection of functional domains in phage fd adsorption protein. J. Mol. Biol. 212, 143.

Strohal, R., Kroemer, G., Wick, G. and Kofler, R. (1987) Complete variable region sequence of a nonfunctionally rearranged kappa light chain transcribed in the nonsecretor P3-X63-Ag8.653 myeloma cell line. Nucleic Acids Res. 15, 2771.

Thomas, M.R. (1994) Simple, effective cleanup of DNA ligation reactions prior to electro-transformation. BioTechniques 16, 988.

Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J. and Johnson, K.S. (1996) Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. Nature Biotechnol. 14, 309.

Weiss, S. and Wu, G.E. (1987) Somatic point mutations in unrearranged immunoglobuline gene segments encoding the variable region of λ light chains. EMBO J. 6, 927.

Wentzell, L.M., Nobbs, T.J. and Halford, S.E. (1995) The SfiI restriction endonuclease makes a four-strand DNA break at two copies of its recognition sequence. J. Mol. Biol. 248, 581.

Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. (1994) Making antibodies by phage display technology. Annu. Rev. Immunol. 12, 433.

Yamanaka, H.I., Kirii, Y. and Ohmoto, H. (1995) An improved phage display antibody cloning system using newly designed PCR primers optimized for Pfu DNA polymerase. J. Biochem. 117, 1218.

Yang, W.-P., Green, K., Pinz-Sweeney, S., Briones, A.T., Burton, D.R. and Barbas III, C.F. (1995) CDR walking mutagenesis for the affinity maturation of a potent human anti-HIV-1 antibody into the picomolar range. J. Mol. Biol. 254, 392.

Zelenetz, A.D. (1992) Construction of complex directional complementary DNA libraries in SfiI. Methods Enzymol. 216, 517.

Zhou, H., Fisher, R.J. and Papas, T.S. (1994) Optimization of primer sequences for mouse scFv repertoire display library construction. Nucleic Acids Res. 22, 888.

Engineering of chicken avidin: a progressive series of reduced charge mutants

Ari T. Marttila[a], Kari J. Airenne[a], Olli H. Laitinen[a], Tikva Kulik[b], Edward A. Bayer[b], Meir Wilchek[b], Markku S. Kulomaa[a],*

[a]Department of Biological and Environmental Science, University of Jyväskylä, P.O. Box 35, 40 100 Jyväskylä, Finland
[b]Department of Biological Chemistry, Weizmann Institute of Science, 76100 Rehovot, Israel

Abstract Avidin, a positively charged egg-white glycoprotein, is a widely used tool in biotechnological applications because of its ability to bind biotin strongly. The high pI of avidin ($\sim$10.5), however, is a hindrance in certain applications due to non-specific (charge-related) binding. Here we report a construction of a series of avidin charge mutants with pIs ranging from 9.4 to 4.7. Rational design of the avidin mutants was based on known crystallographic data together with comparative sequence alignment of avidin, streptavidin and a set of avidin-related genes which occur in the chicken genome. All charge mutants retained the ability to bind biotin tightly according to optical biosensor interaction analysis. In most cases, their thermal stability characteristics were indistinguishable from those of the wild-type avidin. Our results demonstrate that the charge properties of avidin can be modified without disturbing the crucial biotin-binding activity.

© 1998 Federation of European Biochemical Societies.

Key words: Avidin; Protein engineering; Charge mutant; Avidin-biotin technology

1. Introduction

Avidin is a basically charged, tetrameric glycoprotein found in the chicken egg white. Throughout the years, avidin has become a frequently used tool in numerous biotechnological applications, including different localization, diagnostic and separation technologies [1]. Recently avidin has also found its use in affinity-based targeting of drugs and imaging agents with promising results [2–5]. All these applications are generally based on the high affinity (K_d $\sim$10^{-15} M) [6] avidin has for biotin, a low-molecular-weight vitamin, which can be readily attached to biologically active binders and detectable probes. This strong interaction with biotin, combined with the exceptional stability and four biotin-binding sites of avidin (one per subunit) has created the inherent utility and the versatility of the avidin-biotin technology.

Avidin is, however, a positively charged glycoprotein (pI $\sim$10.5) [6], which possesses eight arginine and nine lysine residues [7]. The high pI of avidin and the presence of carbohydrate residues have been a constant hindrance to its use in some applications, due to non-specific binding (mostly charge-related) to extraneous material. For this reason streptavidin, a non-glycosylated and neutrally charged bacterial counterpart of avidin [8], has become the preferred choice in such applications. The preference of streptavidin over egg-white avidin

has prevailed, despite the fact that avidin is more hydrophilic, contains more lysine residues for potential attachment of probes, and is considerably more abundant and cheaper than streptavidin.

In the present work, we wanted to investigate whether the pI of avidin can be reduced using protein engineering, without disturbing significantly the biotin-binding activity or the stability characteristics of avidin. We used sequence comparison of streptavidin [9] and recently cloned avidin-related (*avr*) genes [10], together with the crystallographic structure of avidin [11,12] to design the changes. This approach has allowed us to generate a series of fully functional avidin mutants with pIs ranging from 9.4 down to 4.7. These reduced charge mutants bind biotin in a manner similar to that of wild-type avidin and also display clearly reduced non-specific binding characteristics. Therefore, this study also offers new possibilities for the applications of avidin-biotin technology. Preliminary results of this work were presented at the 8th European Congress on Biotechnology, Budapest 17–21, Hungary, August 1997, Abstracts, p. 158.

2. Materials and methods

2.1. Site-directed mutagenesis and construction of recombinant baculoviruses

Mutagenesis of avidin cDNA [13] was accomplished by the PCR-based megaprimer method [14] using Pfu DNA polymerase (Stratagene, La Jolla, CA, USA). The oligonucleotides for mutagenesis were purchased from either KEBO Lab (Espoo, Finland) or from MedProbe (Oslo, Norway). After digestion with *Bgl*II and *Hind*III (Promega, Madison, WI, USA), the PCR fragments were subcloned into the *Bam*HI/*Hind*III digested pFastBAC1 donor vector (Gibco-BRL, Gaithersburg, MD, USA). The mutations were confirmed by double-stranded sequencing using Sanger's dideoxynucleotide chain termination procedure with an automated DNA sequencer (ALF, Pharmacia Biotech). The recombinant baculoviruses were generated using the Bac-To-Bac baculovirus expression system according to the manufacturer's instructions (Gibco-BRL). The primary virus stocks were amplified for large-scale production of avidin mutants and the titers of the stocks were determined by a plaque assay procedure [15].

2.2. Expression and purification of avidin mutants

Insect cells, Sf9 (ATCC CRL 1711), were maintained as a suspension culture in serum-free Sf-900 II SFM medium (Gibco-BRL) and infected with recombinant baculoviruses at m.o.i. of 0.5–2 pfu/cell. The infection was allowed to proceed for 24 h, after which the culture medium was collected by centrifugation (100×g, 22°C, 5 min) and replaced with fresh biotin-free medium. After the medium change, the infection was continued for another three days. The purification of each avidin mutant was carried out by affinity chromatography using 2-iminobiotin-agarose as previously described by Airenne et al. [16].

2.3. Protein analysis

Electrophoretic analysis was carried out using 15% (w/v) SDS-

*Corresponding author. Fax: (358) (14) 602 221.
E-mail: kulomaa@csc.fi

PAGE with discontinuous buffer system [17]. After electrophoresis, proteins were either stained with Coomassie brilliant blue or blotted onto nitrocellulose membrane for immunostaining according to Airenne et al. [16]. Isoelectric focusing was performed using polyacrylamide gels with pH gradient ranging from 3 to 10. An aliquot of avidin mutants (5 μg) together with the p*I* standards (Bio-Rad) was applied to the gel, and following the run the proteins were visualized by Coomassie staining. The quaternary status of avidin was analyzed by FPLC on a Superose 12 column (Pharmacia) using an LKB HPLC system. A sample (20 μg in 100 μl of phosphate buffer with 0.65 M NaCl, pH 7.2) was applied, and chromatography was carried out at a flow rate of 0.5 ml/min, using the same ionic strength in the equilibration and running buffers. The column was calibrated using bovine γ-globulin, BSA, avidin standard, ovalbumin, ribonuclease and cytochrome *c* as molecular weight markers.

2.4. Interaction analysis of avidin mutants

Binding kinetics were measured using optical biosensor technology (IASyS Manual+, Affinity Sensors, Cambridge, UK). The measurements (R_{eq}, in arc seconds, a measure of the mass on the surface) were carried out using either a commercial biotin cuvette (Affinity Sensors) or by immobilizing 2-iminobiotin onto the carboxymethyldextran cuvette using *N*-hydroxysuccinimide activation. Binding of various concentrations of avidin or avidin mutants onto 2-iminobiotin surface was measured in a 50 mM borate buffer (pH 9.5) containing 1 M NaCl at room temperature. The iminobiotin cuvettes were regenerated with 20 mM HCl. The measurements using the biotin cuvette were carried out using PBS with 1 M NaCl as a binding buffer at room temperature. The kinetic rate constants for association (k_{on}) and dissociation (k_{off}) or the dissociation constant (K_d) were calculated using the Fast Fit program package (Affinity Sensors).

2.5. Thermal stability of avidin mutants

Purified avidin mutants, in the presence or absence of an excess of biotin, were combined with sample buffer (0.125 M Tris-HCl, pH 6.8/ 4% (v/v) SDS/20% (v/v) glycerol/0.004% (v/v) bromophenol blue/10% (v/v) 2-mercaptoethanol) and incubated at selected temperatures, before being subjected to SDS-PAGE as described by Bayer et al. [18]. The gels were stained using Coomassie brilliant blue. The stability of the proteins was followed by dissociation of the tetramer to the monomeric form.

2.6. Non-specific binding assay

Successive dilutions (1 μg, 200 ng, 40 ng and 8 ng) of a DNA sample (salmon-sperm DNA or pGEM plasmid DNA) were applied to nitrocellulose strips. The DNA was fixed to the strips by UV irradiation, after which the strips were quenched using 5× Denhardt's solution. The sample (20 μg in 1 ml PBS) was then added to a strip and incubated at room temperature for 90 min. After that, the strips were washed with PBS/0.05% Tween 20 solution and stained immunochemically as described by Airenne et al. [16].

3. Results

3.1. Design of avidin mutants

In order to reduce the positive charge of avidin, a series of six mutants was constructed with p*I* range from 9.9 down to 4.7, as calculated theoretically from amino acid sequences using the GCG program package (Genetics Computer Group,

Madison, WI, USA). The alterations were done by changing selected basic amino acids to neutral or acidic ones using site-directed mutagenesis (Table 1) and the proteins were named according to the actual p*I*. The selection of amino acids for lowering the p*I* was based on sequence comparison of avidin, streptavidin and avidin-related proteins (AVR), combined with the available crystallographic data of avidin. All the altered amino acids are surface residues and, according to structural information, have no major role in biotin binding or stability of the avidin tetramer. Where possible, we tried to change arginine residues rather than lysines, due to the importance of their terminal amino groups for derivatization of avidin in many applications.

3.2. Expression and purification of avidin mutants

We have shown previously [16] that avidin can be produced efficiently in Sf9 insect cells, using a baculovirus expression system. The expression of the avidin mutants was generally comparable to that of the wild type, although some correlation between production and the number of the mutations was evident – AvdpI9.4 and AvdpI7.9 being best in this respect (data not shown). All the mutants were purified to 95% purity (judged from SDS-PAGE gel) in one step, using affinity chromatography with 2-iminobiotin as a capturing ligand. For successful purification it was essential to use biotin-free culture medium; if biotin is present in the medium, it will block the binding sites of avidin/mutants thus hampering subsequent affinity purification and possible use of the avidin derivative in applied systems.

3.3. Protein chemical analysis

SDS-PAGE analysis showed that the native avidin and charge mutants separated into three components (data not shown), which presumably represented different stages of post-translational modification of glycosylated avidin [16]. There were minor changes in the migration pattern of certain mutants (especially AvdpI4.7), probably due to the extensive charge differences. The isoelectric points of the avidin mutants were determined with isoelectric focusing, using a pH gradient from 3 to 10 (Fig. 1). The experimental results corresponded well with the values theoretically calculated from the amino acid sequences (Table 1). The differences in the case of AvdpI9.4 and AvdpI9.0 are probably due to poor resolution in the upper part of the pH gradient. The experimental value for avidin could not be determined with this system, since its p*I* is over 10, which is out of range of the gel system used in this study. The quaternary status of the charge mutants, in the presence and in the absence of biotin, was examined by FPLC on a Superose 12 column (data not shown). Comparison of the elution profiles with molecular mass standards

Table 1
The physiochemical properties of the avidin charge mutants

Name	Mutations	p*I* calculated	p*I* experimental	K_d (M)[a]
Avd	None	10.4	n.d.	2.0×10^{-8}
AvdpI9.4	R122A, R124A	9.9	9.4	1.4×10^{-7}
AvdpI9.0	R26N, R59A	9.3	9.0	3.1×10^{-8}
AvdpI7.9	R2A, K3E, K9E	8.1	7.9	2.3×10^{-8}
AvdpI7.2	K3E, K9D, R122A, R124A	7.2	7.2	4.5×10^{-7}
AvdpI5.9	R2A, K3E, K9E, R122A, R124A	5.9	5.9	3.1×10^{-7}
AvdpI4.7	R2A, K3E, K9E, R26N, R59A R122A, R124A	4.7	4.7	3.5×10^{-8}

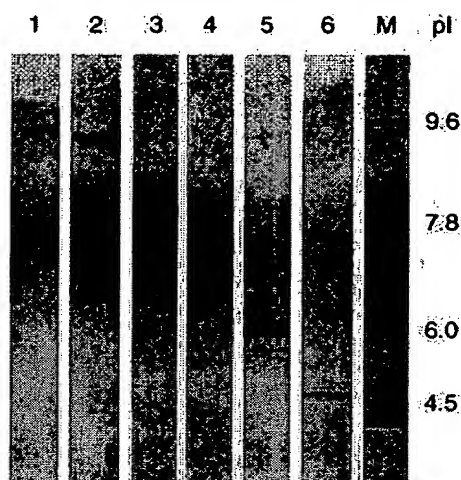[a]Dissociation constant for 2-iminobiotin at pH 9.5; n.d., not determined.

Fig. 1. Isoelectric focusing of avidin charge mutants. Samples (5 μg) of different charge mutants were applied to polyacrylamide gel with pH gradient from 3 to 10. Following the run, proteins were visualized by Coomassie brilliant blue staining. Lanes: 1: AvdpI9.4; 2: AvdpI9.0; 3: AvdpI7.9; 4: AvdpI7.2; 5: AvdpI5.9; 6: AvdpI4.7; M: pI standard.

showed that all the mutants behaved similarly to native avidin (56.6 kDa) forming stable tetramers (the molecular masses of different mutants varied from 55.3 to 60.5 kDa).

3.4. Biotin-binding characteristics of avidin mutants

All six charge mutants bound strongly to 2-iminobiotin, since they were efficiently affinity purified in a single step using this moiety as a capturing ligand. To further characterize the biotin-binding affinities of different mutants, an optical biosensor instrument (IASyS Manual+) was used. When attempts were made to measure the binding constants of mutant proteins to immobilized biotin, no dissociation was observed (data not shown). This indicated that all the mutant proteins exhibited very high affinity constants for biotin – similar to the tenacious binding by the wild-type protein. To obtain precise information concerning the possible differences in affinities among the various mutants, 2-iminobiotin was used as a ligand instead of biotin. Avidin binds 2-iminobiotin at elevated pH with a lower, readily measurable level of affinity. The analysis of the binding curves of the wild-type avidin and different reduced charge mutants, each in several concentrations, were used to determine their dissociation constants to 2-iminobiotin at pH 9.5 (Fig. 2). The calculated $K_d = 2.0 \times 10^{-8}$ M for avidin is in good agreement with that reported previously by Green [6]. The K_d values for different mutants varied from 1.4×10^{-7} M to 3.5×10^{-8} M (Table 1), suggesting that their biotin-binding properties are quite similar to that of wild-type avidin.

3.5. Thermal stability of avidin mutants

In order to determine whether the mutations affected the thermal stability of the avidin tetramer, the avidin mutants and native avidin were diluted in SDS-containing buffer and heated to temperatures between 25°C and 100°C in the absence and presence of biotin. It has been shown previously that under the conditions of this assay biotin stabilizes the quaternary structure of avidin [18]. The transition tempera-

tures for biotin-saturated mutants should therefore be higher than those of the biotin-free derivatives. Biotin-free native avidin starts to dissociate into monomers at temperatures over 57°C (Fig. 3A), whereas with biotin the transition begins only at temperatures near 100°C (Fig. 3B). The results of the thermostability experiment showed that of all the mutant avidins AvdpI4.7 (Fig. 3C and D) was the most stable. Surprisingly, this mutant appeared to be slightly more stable than wild-type avidin. AvdpI9.4 and AvdpI9.0 showed dissociation profiles similar to that of native avidin, but biotin-free AvdpI7.9, AvdpI7.2 and AvdpI5.9 appeared to dissociate, presumably into monomers, at room temperatures when SDS was present (data not shown). Interestingly, in the presence of biotin they behaved in a manner similar to that of native avidin, requiring temperatures near 100°C to dissociate.

3.6. Non-specific binding to DNA

A dot-blot assay was used to investigate the status of the charge mutations with respect to the non-specific binding properties of avidin (data not shown). Wild-type avidin bound strongly to both single- and double-stranded DNA. Most of this binding seems to be charge-related, since there was a correlation between lowering the isoelectric point and reduced binding to DNA; AvdpI4.7 showed the lowest levels of binding to DNA.

4. Discussion

During the last two decades, avidin has evolved into a key component in many biotechnological applications such as affinity-based separations and diagnostic assays [1]. Applications of avidin-biotin technology utilize the extraordinary affinity avidin has for biotin. While the usefulness of avidin is impressive, there are some drawbacks associated with its utilization in some applications. Most of these problems are related to the high pI of avidin. For example the alkaline nature of avidin can cause charge-related binding to DNA and cell surfaces, which may hinder its use in certain circum-
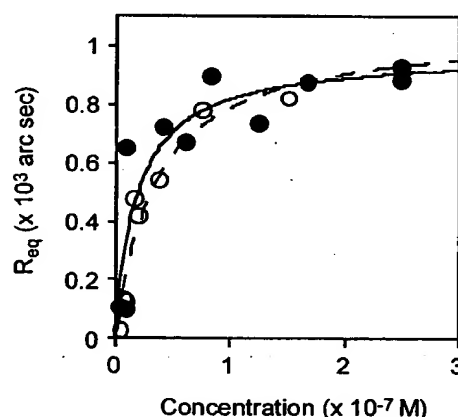
Fig. 2. Interaction of native avidin (closed circles) and AvpI4.7 (open circles) with 2-iminobiotin. Various concentrations of avidin or mutant were added to 2-iminobiotin-coated cuvettes, and binding was measured in pH 9.5 buffer at room temperature using an IASyS optical biosensor. The equilibrium response (R_{eq}) is plotted vs. protein concentration. The K_d of the protein is equal to the concentration at $R_{max}/2$.
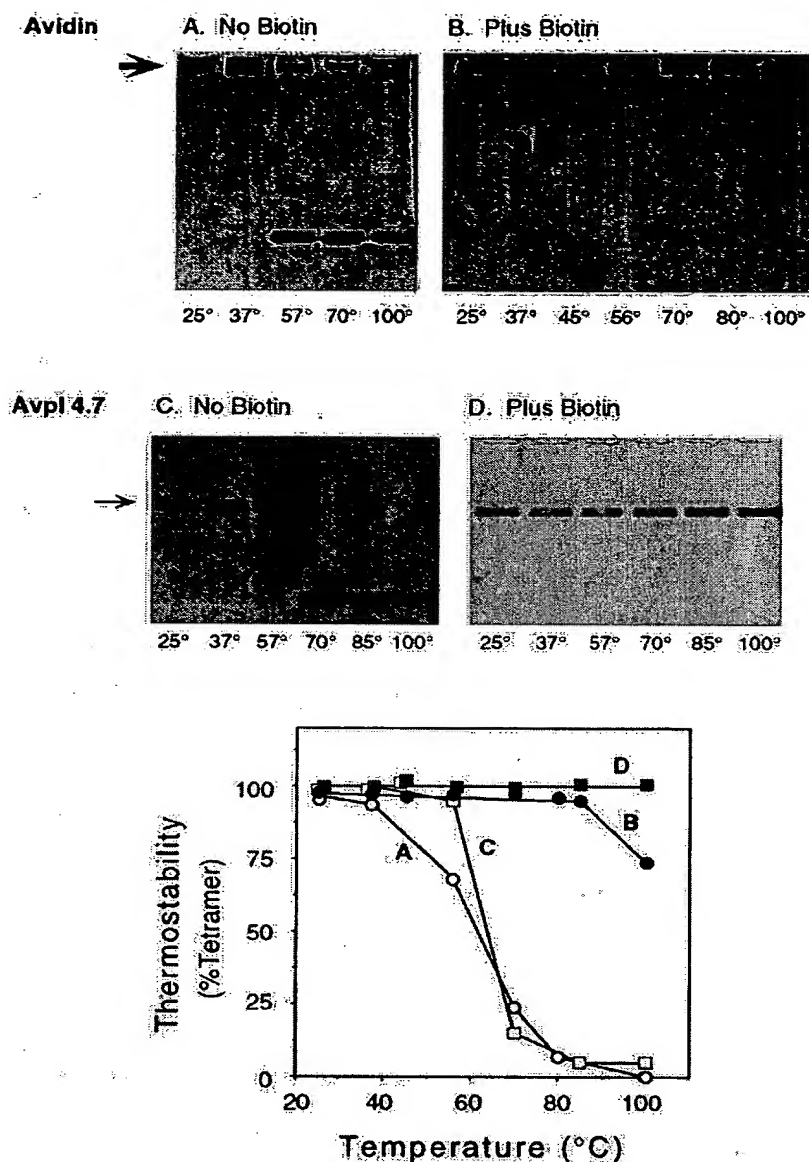
Fig. 3. Temperature-dependent dissociation of native avidin, AvdpI4.7 and their complexes with biotin. Samples of biotin-free and biotin-saturated avidin or AvdpI4.7 were combined with sample buffer and incubated for 20 min at the designated temperatures. The samples were then subjected to SDS-PAGE in 15% separating gels, and the gels were stained using Coomassie brilliant blue. Note that AvdpI4.7 tetramers (C and D, thin arrow) penetrated the upper gel, whereas those of the native avidin (A and B, thick arrow) failed to do so. Densitometry tracings from each of the gels were graphed as a function of temperature: A: avidin; B: avidin+biotin; C: AvdpI4.7; D: AvdpI4.7+biotin.

stances. The high p*I* of avidin has also been a detriment to its use in affinity-based drug targeting, since the positive charge of avidin is considered to be one of the major reasons for its rapid removal from the circulation system [3,19–21].

Site-directed mutagenesis is currently the method of choice to modify properties of a protein or to study the connection between its structure and function. Successful protein engineering requires understanding of the basic concepts of protein biosynthesis and structure. In the case of avidin, the availability of crystallographic data [11,12] together with the sequence information from avidin-related proteins [10] and

streptavidin [9] have made it possible to design such changes in an intelligent manner.

In the present work, we wanted to study whether the high p*I* of avidin can be reduced without losing its stability or high affinity for biotin. For this purpose we constructed a set of reduced charge mutants with isoelectric points ranging from 9.4 down to 4.7. Most of these mutations are based on natural selection and evolution, since they were designed from sequence-based comparison with streptavidin and avidin-related proteins. Therefore Arg-2 was changed to Ala and Lys-3 to Glu according to the sequence of streptavidin. Lys-9 was re-

placed with Glu according to the sequence of the avidin-related protein, AVR1. Arg-26 was modified to Asn according to AVR1 and AVR2, and Arg-59 to Ala, since alanine appears at this position in every AVR protein. In AvdpI7.2, Lys-9 was replaced by Asp instead of Glu due to a PCR error. On the other hand, Arg-122 and Arg-124 were both altered to Ala, because these C-terminal residues were not observed in the crystal structure of avidin [11].

The biotin-binding properties of all charge mutants were similar to those of native avidin, according to the biosensor data. The actual affinities for biotin could not be determined, since it was essentially impossible to remove the bound proteins from the biotin-derivatized cuvette. This was not surprising, since the current biosensor technology can be used to determine affinities only up to $K_a \sim 10^{12}$ M^{-1}. We therefore decided to use 2-iminobiotin as a ligand. Chilkoti et al. [22] have previously demonstrated 2-iminobiotin to be a good reporter for intrinsic streptavidin-biotin interactions and the same is likely to hold true for avidin. The affinities of mutants for 2-iminobiotin (Table 1) were comparable to native avidin, which suggested that their affinities for biotin were also similar (K_as probably in the neighborhood of 10^{13}–10^{15} M^{-1}). This was not so surprising because all the changed amino acids were on the surface of the protein. In this regard, it has been shown previously that, at least in the case of lysozyme, the structural changes resulting from mutations of surface residues are smaller and localized near the mutation, whereas those involving buried residues are larger and may be transmitted to other parts of the protein [23].

As reported earlier [18], native avidin failed to penetrate the separating gel during non-denaturing SDS-PAGE and remained in the aggregated state in the stacking gel (Fig. 3A, arrow). This phenomenon was attributed to the high pI of avidin and possible strong electrostatic interactions with SDS. It is thus interesting to note that, under similar non-denaturing conditions, AvdpI4.7 (Fig. 3C) and most of the other reduced charge mutants migrated in the separating gel in a manner consistent with that of a tetramer.

When the thermal stability of the charge mutants was studied, it was observed that AvdpI9.4, AvdpI9.0 and AvdpI4.7 displayed very similar dissociation profiles to that of the wild-type avidin. In contrast, AvdpI7.9, AvdpI7.2 and AvdpI5.9 showed some differences. In the absence of biotin, the latter derivatives tended to dissociate into monomers already at room temperatures when SDS was present. Interestingly, all of the mutants were stabilized by the presence of biotin; the same mutants required temperatures up to 100°C to undergo denaturation as does wild-type avidin. AvdpI7.9, AvdpI7.2 and AvdpI5.9 all have mutations on Lys-3 and Lys-9 which suggests that either of these lysines (or both) could have some role in the stability of the avidin tetramer. On the other hand, AvdpI4.7 also bears both of these mutations, yet it seemed to be even slightly more stable than wild-type avidin. One possibility is that the additional mutations in AvdpI4.7 (Arg-26 to Asn-26 and Arg-59 to Ala-59) somehow counteract the effect of mutations on Lys-3/Lys-9. The reasons for this enigma remain to be investigated further.

In summary, we have constructed a progressive series of reduced charge mutants of the egg-white protein avidin, with pIs ranging from 9.4 to 4.7. All these mutant proteins bound biotin very strongly, and their thermostability was, in most cases, comparable to that of the wild-type avidin. The mutants also displayed reduced non-specific binding characteristics to DNA when compared to wild-type avidin. Our results demonstrate that the charge properties of avidin can be modified by protein engineering without disturbing the crucial biotin-binding activity. These modified avidins (especially AvdpI4.7) should also prove to be valuable in many applications of avidin-biotin technology, including affinity-based drug targeting and different separation technologies.

References

[1] Wilchek, M. and Bayer, E.A. (1990) Avidin-biotin Technology. Methods Enzymol. 184, Academic Press, San Diego, CA.
[2] Hnatowich, D.J., Virzi, F. and Rusckowski, M. (1987) J. Nucl. Med. 28, 1294–1302.
[3] Pardridge, W.M., Boado, R.J. and Buciak, J.L. (1992) Drug Deliv. 1, 43–50.
[4] Walker, L., Kulomaa, M.S., Bebok, Z., Parker, W.B., Allan, P., Logan, J., Huang, Z., Reynolds, R.C., King, S. and Sorscher, E.J. (1996) J. Drug Target. 4, 41–49.
[5] Shin, S.-U., Wu, D., Ramanathan, R., Pardridge, W.M. and Morrison, S.L. (1997) J. Immunol. 158, 4797–4804.
[6] Green, N.M. (1975) Adv. Protein Chem. 29, 85–133.
[7] DeLange, R.J. and Huang, T.-S. (1971) J. Biol. Chem. 246, 698–709.
[8] Chaiet, L. and Wolf, F.J. (1964) Arch. Biochem. Biophys. 106, 1–5.
[9] Argana, C.E., Kuntz, I.D., Birken, S., Axel, R. and Cantor, C.R. (1986) Nucleic Acids Res. 14, 1871–1882.
[10] Keinänen, R.A., Wallen, M.J., Kristo, P.A., Laukkanen, M.O., Toimela, T.A., Helenius, M.A. and Kulomaa, M.S. (1994) Eur. J. Biochem. 220, 615–621.
[11] Livnah, O., Bayer, E.A., Wilchek, M. and Sussman, J.L. (1993) Proc. Natl. Acad. Sci. USA 90, 5076–5080.
[12] Pugliese, L., Coda, A., Malcovati, M. and Bolognesi, M. (1993) J. Mol. Biol. 231, 698–710.
[13] Gope, M.L., Keinänen, R.A., Kristo, P.A., Conneely, O.M., Beattie, W.G., Zarucki-Schultz, T., O'Malley, B.W. and Kulomaa, M.S. (1987) Nucleic Acids Res. 15, 3595–3606.
[14] Sarkar, G. and Sommer, S.S. (1990) BioTechniques 8, 404–407.
[15] O'Reilly, D.R., Miller, L.K. and Luckov, V.A. (1994) Baculovirus Expression Vectors: A Laboratory Manual, Oxford University Press, New York, NY.
[16] Airenne, K.J., Oker-Blom, C., Marjomäki, V.S., Bayer, E.A., Wilchek, M. and Kulomaa, M.S. (1997) Protein Expr. Purif. 9, 100–108.
[17] Laemmli, U.K. (1970) Nature 227, 680–685.
[18] Bayer, E.A., Ehrlich-Rogozinski, S. and Wilchek, M. (1996) Electrophoresis 17, 1319–1324.
[19] Schechter, B., Silberman, R., Arnon, R. and Wilchek, M. (1990) Eur. J. Biochem. 189, 327–331.
[20] Kang, Y.-S., Saito, Y. and Pardridge, W.M. (1995) J. Drug Target. 3, 159–165.
[21] Rosebrough, S.F. and Hartlet, D.F. (1996) J. Nucl. Med. 37, 1380–1384.
[22] Chilkoti, A., Tan, P.H. and Stayton, P.S. (1995) Proc. Natl. Acad. Sci. USA 92, 1754–1758.
[23] Mathews, B.W. (1995) Adv. Protein Chem. 462, 240–311.

THIS PAGE BLANK (USPTO)

THE BIOTIN-DEPENDENT ENZYMES *

By JOEL MOSS† and M. DANIEL LANE, *Baltimore, Maryland*

CONTENTS

I. Historical Background

The enzymatic role of biotin as a covalently bound "CO_2 carrier" is now unequivocally established. Insight into the mechanism of biotin

action was gained following over a half-century of nutritional, chemical, and metabolic studies, commencing in the latter years of the nineteenth century with the demonstration of the toxic properties of raw egg white (1,2), later shown to be due to avidin (3), a glycoprotein with an extraordinary affinity for biotin (4,5). In 1936, Kögl and Tönnis isolated the yeast growth factor, biotin, from egg yolk in the form of its crystalline methyl ester (6) and determined its empirical formula (7). Biotin was later shown to be identical to vitamin H (8–12), "protective factor X" (13), "the protective factor against egg-white injury" (14), coenzyme R (10,15,16), and certain growth factors (17–19), the former three fractionated on the basis of their ability to prevent avidin-induced biotin deficiency. Du Vigneaud and co-workers determined the correct structure of biotin (20), which was later confirmed by chemical synthesis in the Merck laboratories (21–24), and verified by X-ray crystallography (25–27). By 1950, biotin had been implicated in a number of seemingly unrelated enzymatic processes, including (a) the decarboxylation of oxaloacetate and succinate (28–30), (b) the "Wood–Werkman reaction," that is, the β-carboxylation of pyruvate (31–33), (c) the biosynthesis of aspartate (32,34–44), (d) the biosynthesis of unsaturated fatty acids (39,40,44–53), (e) the deamination of certain amino acids (30,40,54–56), and (f) the synthesis of certain biotin-independent enzymes (57–65). In every instance the enzymatic or metabolic basis for these observations can now be explained in terms of the role of biotin as an enzymatic "CO_2 carrier" or through indirect effects.

One of the early insights into the role of biotin in enzymatic carboxylation was obtained by Wessman and Werkman in 1950 with the oxaloacetate–$H^{14}CO_3^-$ exchange reaction ("Wood–Werkman reaction") catalyzed by cell-free extracts of Micrococcus lysodeikticus (33). Prior exposure of this soluble system to avidin inhibited the exchange reaction. Since treatment of avidin with biotin blocked its inhibitory action, a biotin requirement for the incorporation of bicarbonate into oxaloacetate was indicated. Three years later, Lardy and Peanasky (66) reported that biotin deficiency in rats reduced the ability of liver mitochondrial acetone powder extracts to catalyze the ATP- and divalent cation-dependent carboxylation of propionate to form succinate. Subsequently, Wakil and co-workers showed that highly purified acetyl CoA carboxylase, isolated from avian liver

extracts, was enriched with respect to biotin and sensitive to inhibition by avidin (67–69), suggesting the tight binding of the vitamin to the carboxylase and its probable role in acetyl CoA carboxylation. The earlier isolation and identification of biocytin (ϵ-N-(+)-biotinyl-L-lysine) from yeast extract (70–75) suggested that biotin might be linked to protein by covalent attachment to lysine. In 1962 Kosow and Lane demonstrated that a soluble hepatic enzyme system could be isolated from biotin-deficient rats which catalyzed the covalent attachment of biotin to the lysyl ϵ-amino groups of propionyl CoA apocarboxylase (76–79).

In 1959 Lynen et al. (80) made the interesting observation that an ATP- and Mg^{2+}-dependent carboxylation of free biotin was catalyzed by a bacterial β-methylcrotonyl CoA carboxylase. This model reaction led to the formation of an unstable carboxybiotin derivative which, following its methylation with diazomethane, was identified as 1'-N-carboxybiotin (81,82). The isolation of a carboxylated enzyme intermediate containing stoichiometric quantities of labile carboxyl and covalently bound biotin was accomplished by Kaziro and Ochoa (83) with propionyl CoA carboxylase. Subsequently it was demonstrated with this (84) and several other biotin enzymes (85–89) that the site of carboxylation of the enzyme-bound prosthetic group was the same as that of free biotin.

The proposed role for the enzyme–biotin–CO_2^- intermediate in the reactions catalyzed by biotin enzymes is strongly supported by isotopic exchange experiments (see Section VI), by kinetic data consistent with a "ping-pong" mechanism (see Section V.A) (90–93), and more recently by studies with catalytically active subunits (see Section V.B) (94–100b). The biotinyl prosthetic group can be visualized as a mobile carboxyl carrier migrating between distinct carboxyl donor and acceptor sites on the enzyme. Recent evidence from Vagelos' (94–96) and Lanes' (97–100b) laboratories with the E. coli acetyl CoA carboxylase system shows that these donor (96,97) and acceptor sites (100b), as well as the biotinyl prosthetic group, are localized on different subunits, carboxyl translocation between these sites presumably dependent on the mobility of the prosthetic group.

A large number of biotin-dependent enzymes which carry out diverse types of reactions are known. Many of these enzymes have been obtained in crystalline or homogeneous form and have now been extensively characterized.

II. Types of Biotin-Dependent Enzymes

A. CLASSIFICATION

The biotin enzymes can be classified as carboxylases, transcarboxylases, and decarboxylases, with further subdivision based on substrate specificity (see Table I).

B. REACTIONS*

1. Central Role of Enzyme–Biotin–CO_2^-

The overall reactions catalyzed by biotin-dependent enzymes can be partitioned into two discrete steps (Fig. 1). All of these reactions proceed through a carboxylated intermediate with the carboxybiotinyl prosthetic group functioning as a "mobile carboxyl carrier." The structure of the carboxylated site (Fig. 2) has been verified with four enzymes (Section IV.A), that is, propionyl CoA carboxylase (84),
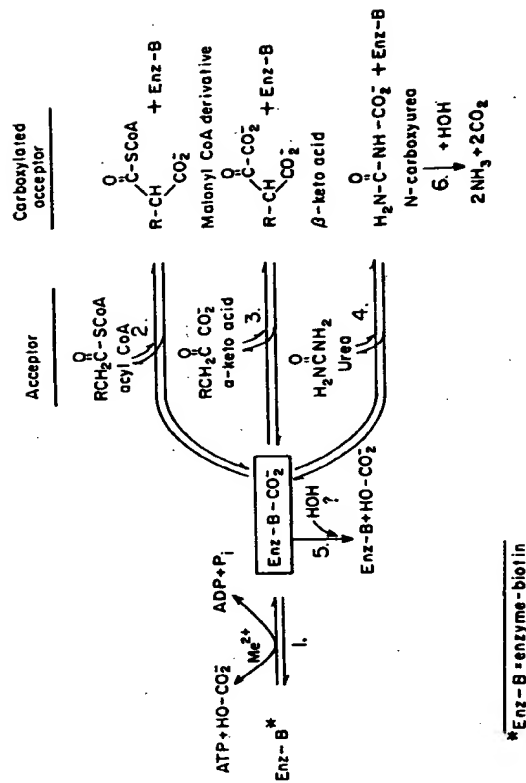
*Enz-B = enzyme-biotin

Fig. 1. Central role of enzyme-biotin-CO_2^- in reactions catalyzed by the biotin-dependent enzymes (reactions 1–6).

* For further details concerning the reactions catalyzed by the biotin-dependent enzymes, the reader should consult Sec. VI.

$$\epsilon\text{-N}\left(\left(1'\text{-N-carboxy-(+)-biotinyl}\right)\text{-L-lysyl enzyme}\right).$$
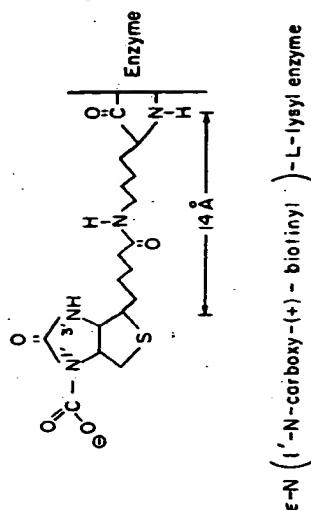
Fig. 2. Structure of the carboxylated biotinyl prosthetic group (ϵ-N-($1'$-N-carboxy-(+)-biotinyl)-L-lysyl enzyme).

β-methylcrotonyl CoA carboxylase (85,86), methylmalonyl CoA-pyruvate transcarboxylase (87), and acetyl CoA carboxylase (88,89), and is probably identical to that of other biotin-containing enzymes.

The initial step involves the formation of carboxybiotinyl enzyme (enzyme-CO_2^-). Carboxylases, which utilize bicarbonate as carboxyl donor, require ATP to drive the formation of the new carbon–nitrogen bond as shown in reaction 1 (Fig. 1). In the case of transcarboxylases and decarboxylases, carboxybiotinyl enzyme formation occurs by an ATP-independent transcarboxylation with either a malonyl CoA derivative (reversal of reaction 2) or a β-keto acid (reversal of reaction 3) serving as carboxyl donor.

In the second step, carboxyl transfer from enzyme-CO_2^- to an appropriate acceptor substrate occurs, the nature of this acceptor being dependent on the specific enzyme involved. With acyl CoA, α-keto acid or amido carboxylases transfer to the position adjacent to the carbonyl group occurs (reactions 2, 3, or 4, respectively), giving rise to a malonyl CoA derivative (101–107), a β-keto acid (108), or N-carboxyurea (109), respectively. Exceptions to this are β-methylcrotonyl CoA (80,110,111) and geranyl CoA (112) carboxylases, where the electrophilic site, hence the site of carboxylation, is displaced to the γ-position by conjugation (Sections VI.C, VI.D). The second step of the transcarboxylase (113,114) reaction makes use of either an α-keto acid (reaction 3) or an acyl CoA derivative (reaction 2) as carboxyl acceptor, thereby completing the transcarboxylation sequence and generating a β-keto acid or malonyl CoA derivative, respectively. With the

---

TABLE I

Classification of the Biotin Enzymes

I. Carboxylases

A. Acyl CoA carboxylases

1. Acetyl CoA carboxylase (acetyl CoA:CO_2 ligase (ADP)), E.C. 6.4.1.2

$$\text{ATP} + \text{HO-CO}_2^- + \text{acetyl CoA} \xrightleftharpoons{\text{Me}^{2+}} \text{malonyl CoA} + \text{ADP} + \text{P}_i$$

2. Propionyl CoA carboxylase (propionyl CoA:CO_2 ligase (ADP)), E.C. 6.4.1.3

$$\text{ATP} + \text{HO-CO}_2^- + \text{propionyl CoA} \xrightleftharpoons{\text{Me}^{2+}}$$
$$\text{(S)-methylmalonyl CoA} + \text{ADP} + \text{P}_i$$

3. β-Methylcrotonyl CoA carboxylase (3-methylcrotonyl CoA:CO_2 ligase (ADP)), E.C. 6.4.1.4

$$\text{ATP} + \text{HO-CO}_2^- + \beta\text{-methylcrotonyl CoA} \xrightleftharpoons{\text{Me}^{2+}}$$
$$\beta\text{-methylglutaconyl CoA} + \text{ADP} + \text{P}_i$$

4. Geranyl CoA carboxylase (geranyl CoA:CO_2 ligase (ADP)), E.C. 6.4.1.—

$$\text{ATP} + \text{HO-CO}_2^- + \text{geranyl CoA} \xrightleftharpoons{\text{Me}^{2+}}$$
$$\text{carboxygeranyl CoA} + \text{ADP} + \text{P}_i$$

B. α-Keto acid carboxylases

1. Pyruvate carboxylase (pyruvate:CO_2 ligase (ADP)), E.C. 6.4.1.1

$$\text{ATP} + \text{HO-CO}_2^- + \text{pyruvate} \xrightleftharpoons{\text{Me}^{2+}} \text{oxaloacetate} + \text{ADP} + \text{P}_i$$

C. Amido carboxylases

1. ATP:urea amidolyase[a]

$$\text{ATP} + \text{HO-CO}_2^- + \text{urea} \xrightleftharpoons{\text{Me}^{2+}}$$
$$\text{ADP} + \text{P}_i + N\text{-carboxyurea} \xrightarrow{+\text{HOH}} 2CO_2 + 2\text{NH}_3$$

II. Transcarboxylases

1. Methylmalonyl CoA:pyruvate transcarboxylase (methylmalonyl CoA:pyruvate carboxyltransferase), E.C. 2.1.3.1

$$\text{(S)-methylmalonyl CoA} + \text{pyruvate} \xrightleftharpoons{}$$
$$\text{oxaloacetate} + \text{propionyl CoA}$$

III. Decarboxylases

A. α-Carboxyacyl CoA decarboxylases

1. Methylmalonyl CoA decarboxylase (methylmalonyl CoA carboxylyase), E.C. 4.1.1.—

$$\text{(S)-methylmalonyl CoA} \rightarrow \text{propionyl CoA} + CO_2$$

B. β-Keto acid decarboxylases

1. Oxaloacetate decarboxylase (oxaloacetate carboxylyase), E.C. 4.1.1.3

$$\text{oxaloacetate} \rightarrow \text{pyruvate} + CO_2$$

[a] Since the mechanism of this reaction has not been established, a systematic name cannot be assigned.

decarboxylases (115–117), it has not yet been established whether the decarboxylation second step (decarboxylation of enzyme-CO_2^-, reaction 5), involves water as the carboxyl acceptor or proceeds via a unimolecular decarboxylation reaction. Nonenzymatic decarboxylation of model N-carboxyimidazolidone derivatives occurs without the intervention of water (Section IV.C.1). (118).

Unlike the first two steps of the amido carboxylase-catalyzed reaction, which are analogous to the 2-step model of the other biotin enzymes, the third step (hydrolysis of N-carboxyurea, reaction 6) does not involve the participation of the biotin prosthetic group (109). with the exception of this reaction and reaction 5, the decarboxylation of "Enz–B–CO_2^-" all of the other partial reactions are readily reversible.

2. Partial Reactions

All of the multistep reactions of the biotin enzymes can be accounted for by appropriate combinations of a few basic types of partial reactions. This is reasonable from the structural point of view, since it appears that biotin enzymes are composed of several nonidentical subunits (94–100b,119–125) and that each component step is probably carried out by a different subunit (Section V.B.) (94–100b). It has been demonstrated with two bacterial enzymes (96,97,121,123) that the biotinyl prosthetic group resides on a small polypeptide chain, distinct from the catalytic subunits, and appears to function in carboxyl translocation between these subunits. This intersubunit oscillation of the carboxylated prosthetic group is supported by kinetic analyses which indicate a "ping-pong" mechanism (Section V.A.) (90–93). Hence the overall reaction is a composite of several partial steps which may be determined by the particular hybridization of subunits in that enzyme complex.

a. Carboxylases. The 2-step reaction sequence shown below (reactions 7a and 7b) is carried out by all biotin-dependent carboxylases. As pointed out previously, the amido carboxylase, that is, urea amidolyase, catalyzes an additional step (reaction 7d) (109) which does not involve the participation of the prosthetic group. Whereas the biotin-dependent steps (reactions 7a–7c), common to all of the carboxylases, are inhibited by avidin, the hydrolysis of the "carboxylated acceptor," N-carboxyurea (reaction 7d) catalyzed by only the amido carboxylase, urea amidolyase, is not. It is possible that the

first partial reaction (reaction 7a), the ATP-dependent carboxylation of the biotinyl group, may involve more than a single chemical step;

$$\text{Enz–biotin} + \text{HO—CO}_2^- + \text{ATP} \xrightleftharpoons{\text{Me}^{2+}} \text{Enz–biotin–CO}_2^- + \text{ADP} + \text{P}_i \quad (7a)$$

$$\text{Enz–biotin–CO}_2^- + \text{acceptor} \rightleftharpoons \text{carboxylated} + \text{Enz–biotin} \quad (7b)$$
$$\text{acceptor}$$

Net:

$$\text{acceptor} + \text{HO—CO}_2^- + \text{ATP} \xrightleftharpoons{\text{Me}^{2+}} \text{carboxylated acceptor} + \text{ADP} + \text{P}_i \quad (7c)$$

acetyl CoA	malonyl CoA
propionyl CoA	methylmalonyl CoA
β-Me-crotonyl CoA	β-Me-glutaconyl CoA
pyruvate	oxaloacetate
urea	N-carboxyurea
geranyl CoA	carboxygeranyl CoA

$$\text{N-carboxyurea} + \text{HOH} \rightarrow 2\text{CO}_2 + 2\text{NH}_3 \quad (7d)$$
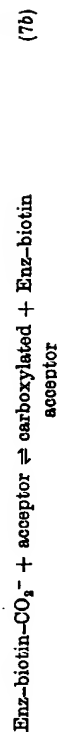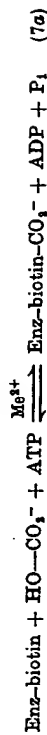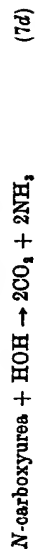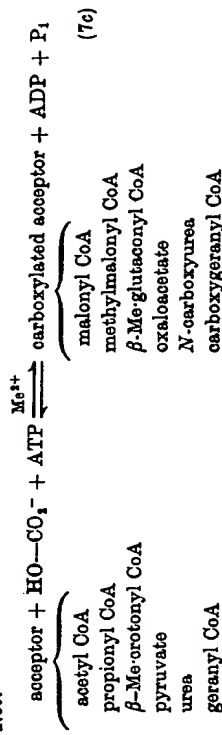
however, there is little evidence to support this notion (Section IV.A). Rather, the occurrence of reaction 7a is supported by the fact that the carboxylases, where investigated, catalyze ATP-^{32}P$_i$ exchange and that maximal rates of exchange are achieved only in the presence of ADP, Me^{2+}, and bicarbonate (Section VI.) (81,125–136). Furthermore, the reciprocal ATP-^{14}C or ^{32}P$_i$-ADP exchange catalyzed by acyl CoA carboxylases exhibits a requirement for P$_i$, Me^{2+}, and bicarbonate (81,126,128,133,134). Neither exchange reaction is dependent on the presence of carboxyl acceptor. With the exception of pyruvate carboxylase (Section VI.E) (132,137), where only ATP-^{32}P$_i$ exchange was found to be avidin-sensitive, these exchange reactions are all inhibited by avidin. In several instances (83–89,128,132,138–144) it has been possible to isolate "substrate amounts" of carboxylated enzyme (enzyme-CO_2^-) free of reactants, following incubation of the enzyme with bicarbonate, ATP, and divalent metal ion or with carboxylated acceptor. Enzyme-CO_2^- can be shown to undergo rapid and quantitative decarboxylation in the presence of ADP, P$_i$, and divalent metal ion (reverse of step 7a) (83,87,128,136,139,145) or quantitative carboxyl transfer to the appropriate acceptor substrate to form carboxylated acceptor (step 7b) (83,86–88,128,132,138–140).

The biotin-containing carboxylases act as weak bicarbonate-dependent ATPases since the carboxylated forms (enzyme-CO_2^-) of

most if not all biotin enzymes undergo slow spontaneous decarboxylation (reaction 8b; $t_{1/2}$ for enzyme-CO_2^- = ~5 to 30 min at pH ~7-8 at ~25°; Section IV.C.1) (83,86,87,139,145,146). Since the enzyme is

$$\text{Enz-biotin} + \text{HO—CO}_2^- + \text{ATP} \underset{}{\overset{Me^{2+}}{\rightleftharpoons}} \text{Enz-biotin—CO}_2^- + \text{ADP} + P_i \qquad (8a)$$

$$\text{Enz-biotin—CO}_2^- \xrightarrow{\text{(slow)}} \text{Enz-biotin} + CO_2 \overset{+\text{HOH}}{\rightleftharpoons} \text{HO—CO}_2^- + H^+ \qquad (8b)$$
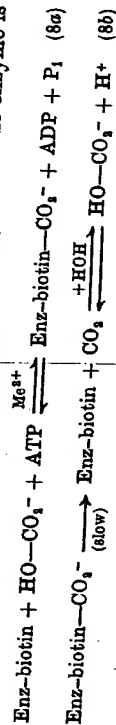
$$\text{Net:} \quad \text{ATP} + \text{HOH} \xrightarrow{\text{HO—CO}_2^-,\ Me^{2+}} \text{ADP} + P_i + H^+ \qquad (8c)$$

rapidly recarboxylated, ATP is cleaved at a rate equal to that of enzyme-CO_2^- decarboxylation. The reaction sequence which accounts for this ATPase activity is shown above. The bicarbonate-dependent ATPase activity can be increased by circumventing the rate-limiting step by arsenolysis (129). Thus the incorporation of arsenate via reaction 9b leads to the formation of ADP-AsO_3^{2-} which spontaneously decomposes to ADP and $HAsO_4^{2-}$ as shown below. Arsenolysis is consistent with the occurrence of the first partial reaction (reaction 9a).

$$\text{Enz-biotin} + \text{HO—CO}_2^- + \text{ATP} \overset{Me^{2+}}{\rightleftharpoons} \text{Enz-biotin—CO}_2^- + \text{ADP} + P_i \qquad (9a)$$

$$\text{Enz-biotin—CO}_2^- + \text{ADP} + HAsO_4^{2-} \overset{Me^{2+}}{\rightleftharpoons} \text{Enz-biotin} + \text{HO—CO}_2^- + \text{ADP—AsO}_3^{2-} + H^+ \qquad (9b)$$

$$\text{ADP—AsO}_3^{2-} + \text{HOH} \rightarrow \text{ADP} + HAsO_4^{2-} + H^+ \qquad (9c)$$
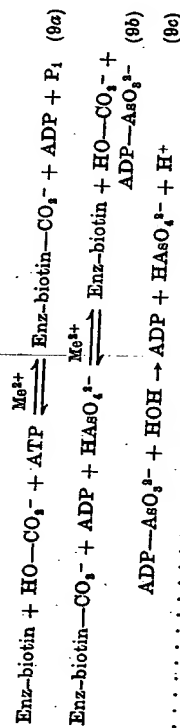
$$\text{Net:} \quad \text{ATP} + \text{HOH} \xrightarrow{\text{HO—CO}_2^-,\ HAsO_4^{2-},\ Me^{2+}} \text{ADP} + P_i + H^+ \qquad (9d)$$

Several lines of evidence support the participation of the second half-reaction (reaction 7b) in the overall carboxylation sequence. First, the direct and stoichiometric transfer of the labile carboxyl group of enzyme-CO_2^- to the appropriate acceptor substrates has been demonstrated with several biotin-dependent carboxylases (83,86-88,128, 132,138-140). Furthermore, most of these carboxylases have been shown to catalyze ^{14}C exchanges between their respective carboxylated acceptor-^{14}C-labeled acceptor pairs (109,125,127-129,131,132,134,136,147-154). The rates of ^{14}C-exchange are compatible with the participation of this step in the overall carboxylation reaction. All of the preceding reactions are inhibited by avidin (Section VII.A) and are not dependent upon bicarbonate, ATP, ADP, P_i or exogenous divalent metal ion. It should

be pointed out that in the case of the manganese-containing α-keto acid carboxylase, pyruvate carboxylase, tightly bound manganese functions only in the second half-reaction (Section IV.C.2) (155).

b. Transcarboxylases. Another type of biotin enzyme which is found in propionibacteria appears to be a functional hybrid of the acyl CoA and α-keto acid carboxylases minus their initial ATP-dependent steps (113,114). This enzyme, methylmalonyl CoA-pyruvate transcarboxylase, while incapable of CO_2 fixation, has the ability to catalyze reversible carboxyl transfer from methylmalonyl CoA to pyruvate through a carboxybiotinyl enzyme intermediate (reactions 10a and 10b).

$$\text{Enz-biotin} + \text{(S)-methylmalonyl CoA} \rightleftharpoons \text{Enz-biotin—CO}_2^- + \text{propionyl CoA} \qquad (10a)$$

$$\text{Enz-biotin—CO}_2^- + \text{pyruvate} \rightleftharpoons \text{Enz-biotin} + \text{oxaloacetate} \qquad (10b)$$

$$\text{Net:} \quad \text{(S)-methylmalonyl CoA} + \text{pyruvate} \rightleftharpoons \text{propionyl CoA} + \text{oxaloacetate} \qquad (10c)$$
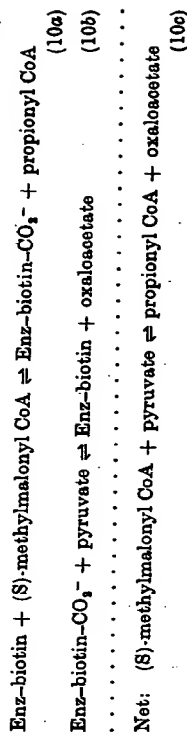
Like the carboxyl transferase components of the acyl CoA and α-keto acid carboxylases, transcarboxylase catalyzes both methylmalonyl CoA-^{14}C-propionyl CoA exchange independent of oxaloacetate and pyruvate, as well as oxaloacetate-^{14}C-pyruvate exchange independent of methylmalonyl CoA and propionyl CoA (91). Similar to other biotin-dependent enzymes which act on an α-keto or β-keto acid substrates, for example, pyruvate carboxylase (155-157) and oxaloacetate decarboxylase (115), transcarboxylase is a zinc- and cobalt-containing metalloenzyme (Section IV.C.2) (122,158).

c. Decarboxylases. The reactions catalyzed by the decarboxylases (115-117) appear to proceed by the 2-step reaction sequence shown below. Because of the rapidity of the decarboxylation of enzyme-biotin-CO_2^- (reaction 11b) has not been possible to isolate the carboxylated

$$\text{Enz-biotin} + \text{carboxylated acceptor} \rightleftharpoons \text{Enz-biotin—CO}_2^- + \text{acceptor} \qquad (11a)$$

$$\text{Enz-biotin—CO}_2^- \rightarrow \text{Enz-biotin} + CO_2 \qquad (11b)$$

$$\text{Net:} \quad \text{carboxylated acceptor} \rightarrow \text{acceptor} + CO_2 \qquad (11c)$$
$$\begin{cases} \text{(S)-methylmalonyl CoA} & \text{Propionyl CoA} \\ \text{oxaloacetate} & \text{pyruvate} \end{cases}$$
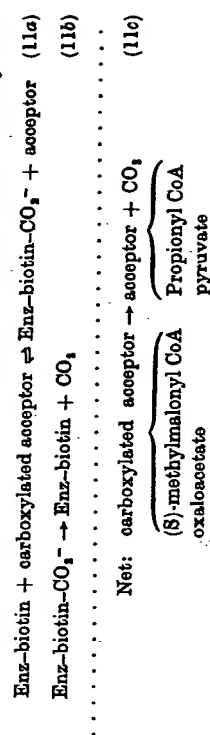
enzyme intermediate. Evidence for the first partial reaction (reaction 11a) has been obtained with methylmalonyl-CoA decarboxylase which

catalyzes an avidin-sensitive exchange between S-methylmalonyl CoA and ^{14}C-propionyl CoA (117).

Similar but abortive reactions are known to occur with carboxylases (Section IV.B.2) (83,86,139,145,146) and transcarboxylase (Section IV.B.2) (87). In this sense, most, if not all biotin-dependent enzymes have the capacity to act as "decarboxylases," albeit at a slower rate. These abortive decarboxylations are distinguished from those catalyzed by the decarboxylases by virtue of the fact that the latter decarboxylation reactions are apparently several orders of magnitude faster than those catalyzed by other biotin-dependent enzymes.

III. The Biotin Prosthetic Group

A. STRUCTURE AND STEREOCHEMISTRY

Following intensive efforts in several laboratories (6,7,20,159–173), the correct structure of biotin was reported in 1942 by du Vigneaud and co-workers (20,166), and subsequently confirmed by chemical synthesis (21–24,174) and X-ray crystallography (25–27,175). Since biotin contains three asymmetric carbon atoms, eight stereoisomers are possible; however, only one, (+)-biotin (see Fig. 3) is biologically active (176). In the active isomer the ureido and tetrahydrothiophene rings are fused cis the aliphatic side chain having a cis configuration with respect to the ureido ring (25). The absolute stereochemistry of (+)-biotin was unequivocally established by X-ray crystallographic analysis of a derivative of the active isomer (bis-p-bromoanilide of 1'-N-carboxybiotin) (26,27).

X-ray crystallographic analysis of (+)-biotin (25) and of the same carboxybiotin derivative described previously (26) revealed that the bicyclic ring system has a boat-like configuration (Fig. 4). The planar ureido ring projects upward at an angle of 62.0° with respect to a plane
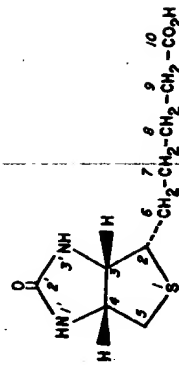
Fig. 3. Absolute configuration of (+)-biotin (2'-keto-3,4-imidazolido-2-tetra-hydrothiophene-n-valeric acid).
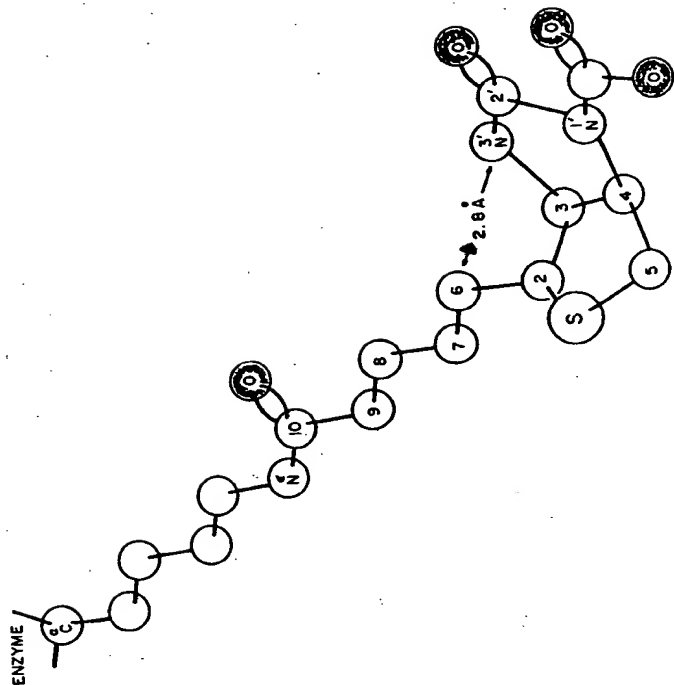
Fig. 4. Diagrammatic representation of ϵ-N-(1'-N-carboxy-(+)-biotinyl)-L-lysyl enzyme (correct absolute configuration).

(plane I) formed by the four carbon atoms of the tetrahydrothiophene ring. Another plane comprising S-1, C-2, and C-5 tilts upward at an angle of 37.8° with respect to plane I (26). Because of the cis orientation of the ureido ring with respect to the aliphatic side chain, C6 resides approximately 2.8 Å from the 3'-N- of the ureido ring (175). Repulsion between the 3'-N and C-6 positions is thought to cause the greater-than-anticipated C_8–C_2–C_8 angle of 119° (175). Steric hindrance due to the juxtaposition of 3'-N and C-6 probably accounts for the favored chemical and (apparent) enzymatic carboxylation at the 1'-N position (82). The X-ray analysis of the carboxybiotin derivative also verified the prior assignment of the 1'-N position as the site of enzymatic carboxylation of free biotin (82) as well as the enzyme-bound biotinyl prosthetic group (Fig. 4) (Section IV.B.1) (84–89).

Early investigations indicated that biotin was probably covalently bound in biological materials to a macromolecular fraction (8,9,177–180). Biotin was found to be covalently attached to propionyl CoA carboxylase through an amide linkage to a lysyl ε-amino group of the enzyme (78). The same mode of attachment has also been demonstrated for several other biotin-dependent enzymes (84–89,133,143,181,182) and probably occurs in all biotinyl proteins. In the case of two bacterial biotin enzymes, *E. coli* acetyl CoA carboxylase and *Propionibacterium shermanii* transcarboxylase, the biotinyl prosthetic group is covalently bound to a low molecular weight carboxyl carrier protein (9000 daltons (97) and 12,000 daltons (121), respectively). Measurements with Corey–Pauling models (Koltun connectors) of ε-*N*-(+)-biotinyl-L-lysine reveals that the maximal distance from the α-carbon of lysine to C-2 of the tetrahydrothiophene ring is 14 A (119). Thus although its covalent attachment firmly anchors the prosthetic group to the apoprotein, the side chain could give the bicyclic ring sufficient freedom of movement to permit translocation between catalytic sites (Section V.B). Additional flexibility may be conferred on the biotinyl group ring system by differential carrier protein interaction with the other subunits.

B. BIOSYNTHESIS

Considerable progress has been made recently in the characterization of the enzymatic steps in biotin biosynthesis. A pathway based on these investigations and consistent with nutritional (183–203), genetic (204–211), and labeling studies (203,212–218) is illustrated in Figure 5.

The synthesis of pimelyl CoA (reactions 1, 2, Fig. 5) from acetyl CoA via malonyl CoA is in agreement with the ^{14}C-labeling pattern of biotin synthesized in *Achromobacter* from 3-hydroxyisovalerate as the principal carbon source, and $H^{14}CO_3^-$ and cysteine-3-^{14}C as labeled precursors (216). Furthermore, the participation of a malonyl CoA intermediate is suggested by the stimulation of biotin biosynthesis by malonate (200,202). However, the enzymatic steps leading to pimelyl CoA have not been characterized, and present evidence does not rule out alternative pathways for its synthesis.

Pimelic acid incorporation into biotin (reaction 3, Fig. 5) was proposed initially on the basis of nutritional (183–188,191,193,195–200, 202) and labeling studies (203,212–214,216). An unidentified amino acid vitamer of biotin, labeled during growth on 1,7-^{14}C-pimelate (203)

and incorporated into biotin (217), was thought to be the condensation product of pimelate and serine or alanine (215). Alternatively it has been proposed (188,216) that cysteine was the 3-C condensing partner, although sulfur incorporation concomitant with a 3-C unit has not been demonstrated and a consistent stimulation of biotin biosynthesis by cysteine has not been obtained (187,188,191,198,202). More recently, Iwahara et al. (219) tentatively identified the vitamer as 7-keto-8-aminopelargonate (7-KAP), and Eisenberg and Maseda (220) confirmed this identification by crystallization of this vitamer from cell cultures. 7-KAP synthetase from *E. coli* is a pyridoxal-dependent enzyme, which may utilize a condensation mechanism similar to that of δ-amino-levulinate synthetase (210,221). This enzyme activity is repressed by growth on biotin (203,210,222).

Fig. 5. Proposed pathway for the biosynthesis of biotin.

The amination step (reaction 4, Fig. 5) by 7,8-diaminopelargonate (DAPA) synthetase has not been characterized. However, the subsequent step, closure of the ureido ring, was demonstrated in whole cell incubations (223), and in cell-free extracts (222,224,225). More recently, a nearly homogeneous preparation of dethiobiotin synthetase has been obtained by Krell and Eisenberg (226). The synthetase has a molecular weight of 42,000 and is composed of 2 weight-homogeneous subunits. Ring closure (reactions 5a, 5b, Fig. 5) requires ATP and CO_2 and gives rise to ADP, P_i, and dethiobiotin. In experiments similar to those of Cooper et al. (227), CO_2 rather than bicarbonate was shown to be the active species. The avidin-insensitive, multistep reaction is thought to be initiated by nucleophilic attack of DAPA on CO_2, forming a carbamate, which is subsequently activated to a carbamyl phosphate derivative by ATP. Nucleophilic attack of the carbamyl phosphate carbon atom by the neighboring amino group eliminating P_i, completes the cyclization (226). Neither dethiobiotin nor biotin is an effective feedback inhibitor (226), although as with KAP synthetase, this enzymatic activity is repressed by biotin (222–224,228).

The conversion of dethiobiotin to biotin (reaction 6, Fig. 5) has been demonstrated in whole cell preparations (194,201). Dethiobiotin can support the growth of yeast (188,190) and is converted by yeast to a biotin vitamer which is utilized by organisms unable to grow on dethiobiotin (192). Labeling studies demonstrated the direct conversion of carbonyl- and carboxyl-labeled-^{14}C-dethiobiotin to biotin in *Aspergillus* (192). The mechanism of sulfur addition has not been determined but probably involves several steps. Dethiobiotin does not repress its own formation (201,230) or that of 7-KAP (230). Furthermore, *E. coli* mutants unable to grow on dethiobiotin accumulate this vitamer (191), and require biotin for growth (191). In intact *E. coli*, biotin represses the conversion of dethiobiotin to biotin (223,229).

In 1944 Gray and Tatum (204) obtained a number of *E. coli* biotin auxotrophs by X-irradiation. More recent studies have defined the nature of the biotin loci, bio A and bio B (204–211,231). The bio A genes can be represented as λ att ABEFGCD (210), where λ att is the lambda phage attachment site; A, the transaminase; B, sulfur incorporation; E, unknown, prior to KAP synthetase; F, G, 7-KAP synthetase; C, unknown, prior to 7-KAP synthetase; and D, ureido ring formation. Studies with λ transducing phages indicate that the bio A gene cluster has a molecular weight of 4.2×10^6, capable of

coding for 10 polypeptide chains of 200 amino acids each (207). The bio B locus is involved in an early step prior to 7-KAP synthetase (209) and is located in the maltose A region (205,231). The several genes coding for steps prior to KAP synthetase may function in pimelate synthesis. Recently, the biotin transport system, the locus of which has not been established, as well as the enzymes involved in biotin synthesis, has been shown to be repressed during growth on biotin (203,210,221,222,229,232–235).

C. ENZYMATIC ATTACHMENT TO THE APOENZYME

1. Mode of Attachment

The terminal step in the synthesis of biotin enzymes appears to be the covalent attachment of the prosthetic group to the apoenzyme to form an active holoenzyme. In biotin-requiring organisms such as the rat, yeast, and certain bacteria, biotin deficiency leads to decreased activities of the biotin enzymes (66,76–79,236–274). These activities can be rapidly restored *in vivo* and *in vitro* by biotin administration or by incubation of biotin-deficient cells (237,266–268,271,274) or cell-free extracts (76,78,79,237,239,241,246,247,259,261,262,270,271) with biotin and the appropriate cofactors. The rapid restoration of holoenzyme activity following biotin administration appears to be independent of new enzyme synthesis, since inhibitors of protein synthesis, for example, actinomycin D, and puromycin, do not block this process (266,268). Furthermore, Jacobs and Majerus (239) have shown that while adipose tissue acetyl CoA carboxylase activity is markedly depressed by biotin deficiency, the level of immunologically reactive carboxylase protein does not decrease and, more importantly, does not increase following restoration of carboxylase activity after biotin administration. Hence new enzyme synthesis does not appear to be involved in the rapid restoration process. Therefore, although biotin deficiency leads to decreased holoenzyme levels, there is a concomitant rise in apoenzyme levels so that the total enzyme (apo- plus holo-) concentration remains constant or actually increases (76,259,261,267,270,271).

It has been demonstrated with cell-free systems (76–79,238,239,246, 249,253,256,257,259–261) that biotin-deficient organisms contain specific synthetases which catalyze the covalent attachment of biotin to the apoenzyme to form active holoenzyme. The first evidence for a synthetase was obtained by Kosow and Lane (76–79) with a soluble
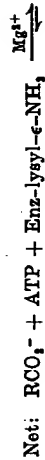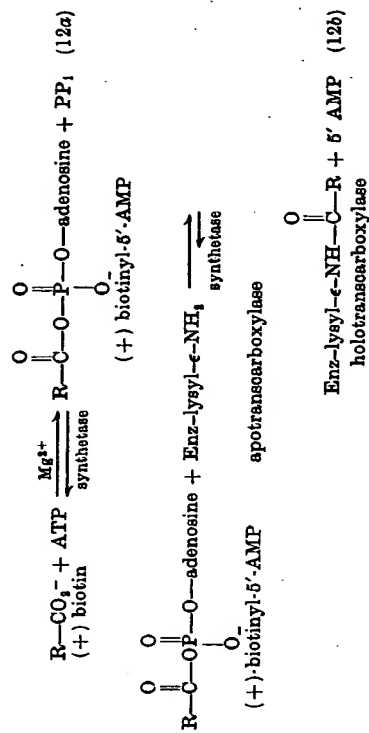
propionyl CoA holocarboxylase-synthesizing system obtained from biotin-deficient rat liver. It was found that in addition to (+)-biotin, both ATP and a divalent metal ion, such as Mg^{2+}, were required for holocarboxylase synthesis. The same system was found (78,79) to catalyze the ATP- and Mg^{2+}-dependent covalent attachment of ^{14}C-biotin to protein and provided a means for determining the nature of the attachment of the prosthetic group to the apoenzyme. The earlier isolation of biocytin (ϵ-N-biotinyl-L-lysine) from yeast extract by Wright et al. (73) suggested that the site of attachment might be a lysyl residue on the enzyme. Following the incubation of ^{14}C-biotin with the propionyl holocarboxylase-synthesizing system, the ^{14}C-holocarboxylase was partially purified and then subjected to proteolytic digestion with Pronase (Streptomyces griseus protease) (275,276). The sole ^{14}C-labeled derivative in the Pronase digests was isolated and subsequently identified as biocytin (78). It was evident, therefore, that the biotinyl prosthetic group of propionyl CoA carboxylase was covalently bonded to a lysyl ϵ-amino group of the apoenzyme as illustrated in Figure 2. This type of linkage was subsequently detected in propionyl CoA carboxylase from other sources (84,181,182) and several other biotin enzymes, that is, β-methylcrotonyl CoA carboxylase (85,86), transcarboxylase (87,181,277), and acetyl CoA carboxylase (88,89,143). This mode of attachment extends also to lipoic acid, which is linked through its carboxyl group in amide linkage to lysyl ϵ-amino groups of pyruvate dehydrogenase (278,279).

2. Partial Reactions

The resolution of several holoenzyme-synthesizing systems (propionyl CoA carboxylase (78,79,246,260), transcarboxylase (254,255), and β-methylcrotonyl carboxylase (259)) from biotin-deficient organisms into two essential protein components, that is, apoenzyme and specific synthetase (biotin-apoenzyme ligase (AMP)), made possible the study of the reaction sequence and mechanism. Lane et al. (254,255) concluded on the basis of studies with extensively purified preparations of the synthetase and methylmalonyl CoA-pyruvate apotranscarboxylase from Propionibacterium shermanii that the synthetase-catalyzed reaction involves at least two steps (reactions 12a and b).

The occurrence of the first partial reaction (reaction 12a) is supported by the observation (255,260) that the synthetase catalyzes a biotin-dependent ATP-$^{32}P^{32}P_i$ exchange. This exchange reaction was found

to be relatively specific for (+)-biotin in that neither (−)-biotin, valerate, nor n-caprylate can replace (+)-biotin (255). O-Heterobiotin,

$$R-CO_2^- + ATP \xrightarrow[\text{synthetase}]{Mg^{2+}} R-\overset{O}{\overset{\|}{C}}-O-\overset{O}{\overset{\|}{P}}-O-\text{adenosine} + PP_i \quad (12a)$$
$$\text{(+) biotin} \qquad\qquad \overset{|}{\underset{O_-}{}}$$
$$\text{(+) biotinyl-5'-AMP}$$

$$R-\overset{O}{\overset{\|}{C}}-O\overset{O}{\overset{\|}{P}}-O-\text{adenosine} + \text{Enz-lysyl-}\epsilon\text{-NH}_2 \xrightarrow{\text{apotranscarboxylase}}$$
$$\overset{|}{\underset{O_-}{}}$$
$$\text{(+)-biotinyl-5'-AMP}$$

$$\text{Enz-lysyl-}\epsilon\text{-NH}-\overset{O}{\overset{\|}{C}}-R + 5' \text{ AMP} \quad (12b)$$
$$\text{holotranscarboxylase}$$

Net: $RCO_2^- + ATP + \text{Enz-lysyl-}\epsilon\text{-NH}_2 \underset{}{\overset{Mg^{2+}}{\rightleftharpoons}} \text{Enz-lysyl-}\epsilon\text{-NH}-\overset{O}{\overset{\|}{C}}R + AMP + PP_i \quad (12c)$

in which oxygen is substituted for sulfur in the tetrahydrothiophene ring, supports an exchange rate 6% that of biotin. Similar findings were subsequently made by McAllister and Coon (260) with the corresponding synthetase for the propionyl CoA apocarboxylase from rabbit liver. However, it was found also that (+)-biotin sulfone, (+)-norbiotin, dethiobiotin, homobiotin, and 2-n-valeric acid-3,4-diaminotetrahydrothiophene did not support the exchange reaction. The rates of ATP-$^{32}P^{32}P_i$ exchange catalyzed by both synthetases are significantly greater than the rate of holoenzyme synthesis (255,280) which is compatible with the participation of the first partial reaction in the overall synthetic process. The low K_m value for (+)-biotin in holotranscarboxylase synthesis ($K_m = 4 \times 10^{-7}$ M) approximates that for (+)-biotin in the ATP-$^{32}P^{32}P_i$ exchange reaction ($K_m = 5 \times 10^{-7}$ M) (255), which suggests a common role for the synthetase in both processes. These low K_m values are compatible with biotin being a natural substrate. Similar observations were made subsequently with the liver holocarboxylase-synthesizing system (256,260), although somewhat lower K_m values ($K_m = 3.5 \times 10^{-9}$ M) were obtained. Catalysis of the biotin-dependent ATP-$^{32}P^{32}P_i$ exchange by synthetases

TABLE II
Specificity of Holoenzyme Synthetases toward various apoenzymes[a]

Synthetase preparation	Specific activity[b]			
	Rat-liver propionyl CoA apocarboxylase	Bacterial β-methyl-crotonyl CoA apocarboxylase	Bacterial apotrans-carboxylase	Yeast acetyl CoA apo-carboxylase
Rabbit liver	0.4 (+)[c,d]	3.1 (+)[f]	0	
Yeast	1.6	36.7	41.4 (+)[h]	
Propionibacterium shermanii	2.6 (+)[e]	0.2 (+)[g]	28.9	(+)[c,h]

a Most of the data in this table were taken from a paper by McAllister and Coon (260).
b Expressed in units of holoenzyme formed per minute per milligram of protein of the synthetase preparation.
c Values in parentheses indicate that successful holoenzyme synthesis has been accomplished by another investigator although a quantitative comparison is not possible.
d References 256,257.
e Reference 79.
f Reference 258.
g Reference 259.
h Reference 238.

from several organisms (255,260) was reminiscent of fatty acid (280–286) and amino acid (287–289) activation and suggested that enzymatic activation of the side-chain carboxyl group might give rise to the corresponding biotinyl adenylate.

(+)-Biotinyl-5'-AMP was chemically synthesized (255,290) from (+)-biotin and 5'-AMP by the dicyclohexylcarbodiimide method (291) and found to completely replace (+)-biotin, ATP, and Mg^{2+} for holotranscarboxylase (255) and holocarboxylase (245,249,257,259,260) synthesis. The rate of holoenzyme formation with (+)-biotinyl-5'-AMP is comparable to (255) the rate of synthesis with ATP, Mg^{2+}, and (+)-biotin. Furthermore, the rate of synthetase-catalyzed ATP-β,γ-^{32}P formation (reverse of reaction 12a) from the biotinyl adenylate and $^{32}P^{32}P_i$ is comparable to that of ATP–$^{32}P^{32}P_i$ exchange (255). Supporting evidence for an ATP-dependent activation of the biotin side-chain carboxyl group was obtained by Höpner and Knappe (259), who found that biotinyl hydroxamate is formed when biotin, ATP, and Mg^{2+} are incubated in the presence of high concentrations of hydroxylamine.

3. Specificity

In the synthesis of *active* holoenzyme, specificity for biotin analogues is determined both by the synthetase and by the structure of the biotinyl acceptor site on the apoenzyme. The synthetase may catalyze the acylation of the correct lysyl ε-NH_2 group by a biotin analogue, but an incompetent holoenzyme may be formed. Therefore the specificity of the synthetase for biotin analogues is best assessed using the ATP–$^{32}P^{32}P_i$ exchange reaction. As pointed out earlier, the animal and bacterial synthetases exhibit a high degree of specificity toward biotin analogues when tested by the exchange assay; only O-heterobiotin supports ATP–$^{32}P^{32}P_i$ exchange (255). Although O-heterobiotin cannot replace (+)-biotin for active holoenzyme synthesis, it is a competitive inhibitor with respect to (+)-biotin, which indicates that it does bind to the synthetase. Furthermore, it has been demonstrated (254) that the *P. shermanii* synthetase catalyzed the attachment of this analogue to the homologous apotranscarboxylase; however, the holotranscarboxylase analogue formed is enzymatically inactive. This is interesting since free O-heterobiotin can bind at the catalytic sites of several biotin-containing carboxylases as evidenced by the fact that the free analogue is enzymatically carboxylated (Section IV.B.1). It is

possible that the stringent steric requirements of the site(s) on the enzyme, which interact with the bicyclic ring, cannot be met once the analogue becomes covalently attached to the apoenzyme.

The specificity requirements of the synthetases for acceptor apoenzyme are not as rigorous as for biotin analogues. The results from several laboratories show (Table II) that synthetases from different species can catalyze the acylation by biotin of apoenzymes of various types and from different sources. However, these synthetases are highly selective in that biotin is attached to specific lysyl ε-NH_2 groups at or near the potential catalytic sites of the apoenzymes. The broad specificity of the synthetases toward apoenzymes from different sources indicates that the structures of these enzymes in the vicinity of the biotin attachment site are probably similar.

4. Regulation

A novel mode of regulation of the synthesis of pyruvate holo-carboxylase has been demonstrated in *Bacillus stearothermophilus* by Kornberg and co-workers (261,270). The formation of the holoenzyme from its apoenzyme catalyzed by cell-free extracts is dependent on acetyl CoA, an allosteric activator of pyruvate carboxylase, as well as biotin, ATP, and Mg^{2+}. Holoenzyme synthesis from (+)-biotinyl-5'-AMP and the apoenzyme in the absence of biotin, ATP, and Mg^{2+} is also activated by acetyl CoA. L-Aspartate, a negative allosteric effector of pyruvate carboxylase in this organism, inhibits holo-carboxylase synthesis by antagonizing the action of acetyl CoA (261). The sigmoidal character of the kinetic responses to acetyl CoA and L-aspartate indicates cooperative interactions of both of these ligands with an enzymatic component of the system. Since the apoenzyme would be expected to have binding sites for both ligands, it appears likely that these kinetic effects result from conformational changes in this component, rather than in the synthetase. In an independent study with the avian liver enzyme, Madappally and Mistry (271) found that acetyl CoA did not activate holoenzyme synthesis from apocarboxylase and biotin.

Recent evidence (267) indicates that the synthesis of the apoenzyme may be controlled by the biotin status of the organism. Birnbaum (267) found that the total quantity of acetyl CoA carboxylase (apoenzyme plus holoenzyme) increased approximately fivefold because of biotin deficiency in *Lactobacillus plantarum*. Whereas in biotin-deficient cells >95% of the carboxylase was in the apoenzyme form, in biotin-sufficient cells holoenzyme accounted for essentially all of the carboxylase. It is evident that the synthesis of the apoenzyme is regulated either by biotin per se, that is, by "coenzyme repression" (267), or through an indirect effect reflecting the biotin status of the organism.

IV. Mechanistic Considerations

The initial event in biotin-dependent reactions is the formation of carboxybiotin through nucleophilic attack of the ureido ring (Section IV.B.1) on an electrophilic center of a carboxyl donor, that is, bicarbonate ($HO-CO_2^-$) or an activated carboxyl group. Subsequent carboxyl transfer from carboxybiotin necessitates electrophilic activation of the carboxyl carbon atom, as well as enzyme-mediated nucleophilic activation of the acceptor molecule. Similarly, decarboxylation requires an analogous electrophilic activation of carboxybiotin.

A. "CO2"

Under physiological conditions "CO_2" exists both as dissolved CO_2 and its hydrated forms (H_2CO_3, HCO_3^-, and CO_3^{2-}), the dominant species being HCO_3^- and CO_2. Enzymatic carboxylation reactions are known in which either CO_2 (226,227,295-297) or HCO_3^- (129,292-297)* is the active species. Chemical considerations suggest that CO_2 is a better electrophile than HCO_3^- and therefore a superior carboxylating species. However, the anionic character of HCO_3^- makes it a more attractive ligand for interaction at the active site of the enzyme (298). CO_2 is a linear symmetrical molecule (299,300) in which delocalization of electrons toward the two oxygen atoms renders the carbon atom particularly susceptible to nucleophilic attack. In contrast, bicarbonate because of resonance stabilization is apparently less prone to nucleophilic attack and, therefore, might be expected to require enzyme-mediated electrophilic activation. Precedent for the electrophilic activation of bicarbonate has been obtained with magnesium methyl carbonate, an analogue of bicarbonate (301-306). Stiles (301, 302) has demonstrated that this reagent acts as an electrophilic carboxylating species in reactions with ketones which contain enolizable methyl or methylene groups. The electrophilic character of the carbonate carbon atom is no doubt enhanced by Mg^{2+}, thereby increasing its susceptibility to nucleophilic attack. Hydration of CO_2 is accompanied by substantial angular distortion with reduction of the O-C-O bond angle from 180° to about 120° (300). The bond angles of a carboxylate anion approximate those of bicarbonate. Thus carboxylations involving CO_2 require a change in hybridization of the carbon orbitals from sp to sp^2, whereas carboxylations involving HCO_3^- do not. The slow rate of hydration of CO_2 has been attributed to $sp \rightarrow sp^2$ rehybridization (300) and suggests that a similar kinetic effect may apply in carboxylation reactions requiring CO_2 as the active species. The analogous hydration of SO_2 which is not accompanied by any significant angular distortion is 8 orders of magnitude faster than the hydration of CO_2 (300,307).

* It has not been possible to discriminate between HCO_3^-, H_2CO_3, and CO_3^{2-} although under the usual experimental conditions HCO_3^- is the most abundant species.

An understanding of the mechanistic details of an enzymatic carboxylation reaction requires the identification of the active species of "CO$_2$" which participates in the reaction mechanism. Two different approaches have been utilized to answer this question in the case of the biotin-dependent carboxylases. It was demonstrated by Kaziro et al. (129) in experiments with ^{18}O-labeled bicarbonate that, in the course of propionyl CoA carboxylation, one bicarbonate oxygen atom was incorporated into orthophosphate for every two bicarbonate oxygen atoms incorporated into the carboxyl group of methylmalonyl CoA. In view of these results, as well as the isotopic exchange studies discussed previously (Section II.B.2a), a concerted reaction between enzyme-bound biotin, HCO$_3^-$, and the γ-phosphoryl group of ATP was proposed (129) as shown in Figure 6. The bicarbonate anion is visualized as a nucleophilic reagent which attacks the γ-phosphoryl phosphorus atom displacing ADP in concert with the nucleophilic attack by the 1'-N of biotin on bicarbonate. An analogous mechanism has been suggested (293) based on similar ^{18}O-incorporation experiments in the phosphoenolpyruvate carboxylase-catalyzed reaction.

Another approach for determining the active species of "CO$_2$" in enzymatic processes is based on the fact that equilibrium of the CO$_2 \rightleftharpoons$ H$_2$CO$_3 \rightleftharpoons$ HCO$_3^-$ reaction is not reached for more than 1 min at temperatures below 15°. When the initial reactants are CO$_2$ and H$_2$O (295,295a,308). Since the generation of HCO$_3^-$ (or attainment of ^{14}C-isotopic equilibrium) can be made rate-limiting, the rate of carboxylation (or rate of fixation of ^{14}C-activity) is dependent on

ATP + HCO$_3^-$ + ENZ–BIOTIN

$\xrightarrow{\text{Mg}^{2+}}$

^{18}O$^-$ + ADP + P$_i$(^{18}O)
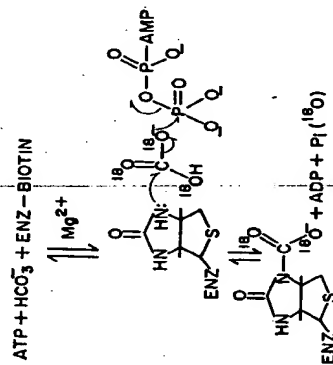
Fig. 6. Possible concerted mechanism for the carboxylation of biotin (from Kaziro et al. (129)).

1. ATP + E-biotin $\rightleftharpoons$ E$\diagdown$P$\diagup$(ADP) $\rightleftharpoons$ ADP + E$\diagdown$P$\diagup$biotin

   $\Updownarrow$ +HCO$_3^-$

   (P$_i$)E(ADP)$\diagup$biotin$\diagdown$CO$_2^-$ $\rightleftharpoons$ E-biotin~CO$_2^-$ + ADP + P$_i$

2. E-biotin~CO$_2^-$ + Acceptor $\rightleftharpoons$ Carboxylated Acceptor + E-biotin
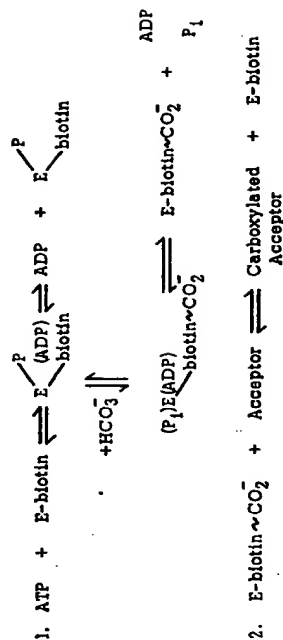
Fig. 7. Alternative stepwise mechanism for the carboxylation of biotin.

(a) whether CO$_2$ or HCO$_3^-$ is the active species and (b) whether CO$_2$ or HCO$_3^-$ is bound first by the enzyme. Investigations based on this approach indicated that HCO$_3^-$ is the active species of "CO$_2$" utilized by pyruvate carboxylase (295–297). Although these approaches do not prove conclusively that HCO$_3^-$ per se participates in the mechanism, in combination they show that HCO$_3^-$ must bind initially and that if an –OH is abstracted at the active site to produce bound CO$_2$, then (a) the –OH abstracted is transferred directly to the γ-phosphoryl phosphorus atom, and (b) the newly formed CO$_2$ does not exchange with CO$_2$ in the medium at a rate faster than carboxylation. However, since both techniques yielded the same result with two different biotin-dependent carboxylases, HCO$_3^-$ is most likely the active species.

Although the concerted mechanism for the formation of carboxy-biotin by acyl CoA carboxylases (Fig. 6) is consistent with the dependence of ATP–^{32}P$_i$ exchange on HCO$_3^-$ and ADP and the dependence of ATP–^{14}C-ADP exchange on HCO$_3^-$ and P$_i$ (81,125–136), a stepwise mechanism cannot be excluded. Indeed, the observation by Scrutton and Utter (137) that pyruvate carboxylase, unlike the acyl CoA carboxylases, catalyzes a slow P$_i$- and HCO$_3^-$-independent ATP-^{14}C-ADP exchange reaction is difficult to rationalize in terms of the concerted mechanism. Since this exchange reaction is not inhibited by avidin, the participation of the biotin prosthetic group appears unlikely. These results suggest that either a more complex carboxylation mechanism or an abortive pathway is involved. A mechanism similar to one suggested by Scrutton and Utter (137) and involving a phosphorylated enzyme intermediate is shown in Figure 7. Initial phosphorylation of the enzyme by ATP may be followed by either the

release of ADP to form phosphoryl-enzyme via the abortive reaction (Fig. 7, upper right) or reaction with HCO_3^- to form carboxy-enzyme-(P_i) (ADP). The former route would account for the observed slow rate of carboxylase-catalyzed $ATP-^{14}C\cdot ADP$ exchange which is independent of P_i or HCO_3^- (137). The dependence of oxaloacetate—$H^{14}CO_3^-$ exchange on ADP and P_i (137) rules out the possible conversion of the abortive product, phosphoryl-enzyme, to carboxy-enzyme except in the presence of ADP and HCO_3^-. Furthermore, it has been demonstrated with acetyl CoA carboxylase that decarboxylation of enzyme-CO_2^- does not occur in the presence of P_i (309), which indicates that, if a phosphoryl-enzyme intermediate is involved, ADP is essential for its formation from enzyme-CO_2^-. It is possible that the reactions of biotin-dependent carboxylases proceed through a phosphoryl-enzyme intermediate and that the productive pathway involves ordered release of P_i and ADP, either subsequent to or simultaneous with carboxybiotin formation.

An equally plausible stepwise mechanism consistent with most of the preceding data involves the initial formation of tightly bound carbonyl phosphate (reaction 13b) (129). In the second step (reaction 13b),

$$\text{Enz–biotin} + \text{HO—CO}_2^- + \text{ATP} \underset{\text{Me}^{2+}}{\rightleftharpoons} \underset{\text{(ADP)}}{\text{Enz–biotin}} \quad (13a)$$

$$\begin{array}{c}\text{O}\\\parallel\\(\text{HO—C—OPO}_3{}^{2-})\end{array}$$

$$\underset{\text{(ADP)}}{\text{Enz–biotin}} \begin{array}{c}\text{O}\\\parallel\\(\text{HO—C—OPO}_3{}^{2-})\end{array} \xrightarrow{\text{Me}^{2+}(?)} \text{Enz–biotin–CO}_2^- + \text{ADP} + P_i \quad (13b)$$

$$\text{Enz–biotin–CO}_2^- + \text{acceptor} \rightleftharpoons \text{Enz–biotin} + \underset{\text{acceptor}}{\text{carboxylated}} \quad (13c)$$

$$\text{Net: HO—CO}_2^- + \text{ATP} + \text{acceptor} \underset{\text{Me}^{2+}}{\rightleftharpoons} \underset{\text{acceptor}}{\text{carboxylated}} + \text{ADP} + P_i \quad (13d)$$
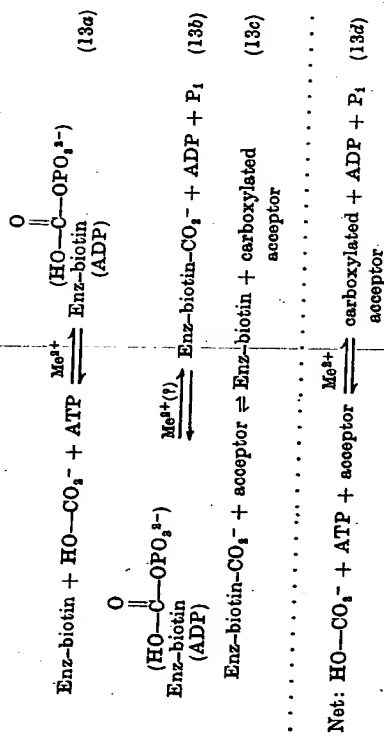
carbonyl phosphate would carboxylate the biotin prosthetic group with release of P_i and bound ADP. As suggested by Griffith and Stiles (310), coordination complexes of phosphoric-carbonic anhydride, of which carbonyl phosphate is representative, are potential carboxylating agents. In order to be consistent with the exchange studies, ADP release from the enzyme could not precede the carboxylation of biotin.

A similar ordered release of ADP was proposed in the previous mechanism (Fig. 7) in order to be compatible with the isotopic exchange studies.

B. BIOTIN

1. Carboxylation

In 1959 Lynen et al. (80) made the interesting observation that bacterial β-methylcrotonyl CoA carboxylase catalyzes the carboxylation of free (+)-biotin in the absence of its natural substrate. The carboxylated product was found to be relatively unstable, particularly at acid pH, and therefore was converted with diazomethane to its dimethyl ester prior to isolation (Fig. 8a). This derivative was found to be identical with the isomer of the dimethyl ester of N-carboxybiotin formed predominantly by chemical synthesis (Fig. 8c) (81,82). It was concluded that the 1'-N rather than the 3'-N position was the site of carboxylation, since the X-ray crystallographic studies of Traub (25, 175) indicated probable steric hindrance of the 3'-N position by C-6 of the aliphatic side chain (Section III.A and Fig. 4). It was conclusively
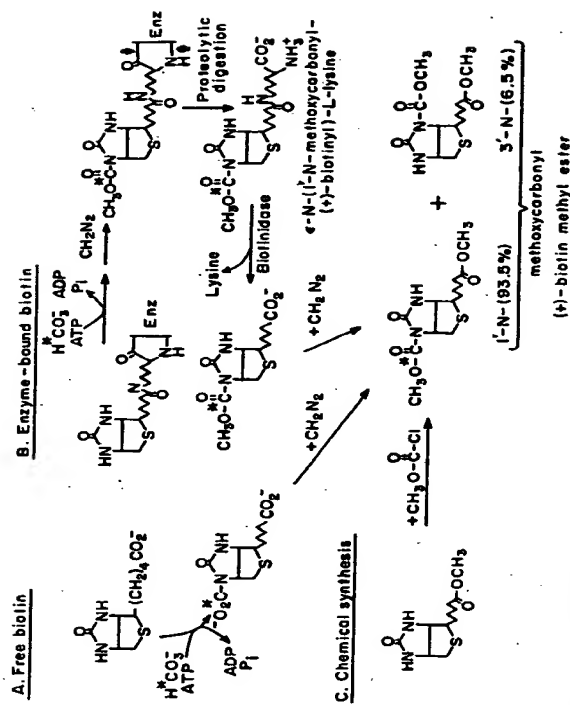

Fig. 8. Identification of the site of carboxylation of free and enzyme-bound biotin.

demonstrated by X-ray crystallographic analysis (26) that the methoxy-carbonyl group of the methylated carboxylation product was located at the 1'-N position. The site of carboxylation on the covalently bound prosthetic group was confirmed subsequently with β-methylcrotonyl CoA carboxylase (85,86), propionyl CoA carboxylase (84), trans-carboxylase (87), and acetyl CoA carboxylase* (88,89). The basic procedure followed is outlined in Figure 8b. Following carboxylation of the enzyme either with ATP, $H^{14}CO_3^-$, and Mg^{2+} or in the case of transcarboxylase with methylmalonyl-3-^{14}C-CoA, enzyme-$^{14}CO_2^-$ was resolved from reactants by gel filtration. Stabilization of the labile carboxyl group of enzyme-$^{14}CO_2^-$ was accomplished by methylation with diazomethane, the methylated protein subjected subsequently to digestion with proteolytic enzymes. The principal radioactive product in the proteolytic digests was identified as ϵ-N-(1'-N-carbomethoxy-(+)-biotinyl)-L-lysine by comparison of its chromatographic properties with the chemically synthesized compound and by further degradation to 1'-N-carbomethoxybiotin with biotinidase.
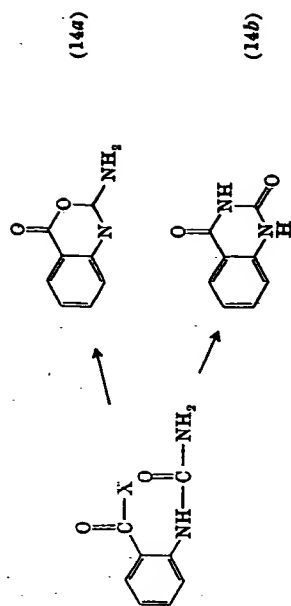
It was thought for some time that the ability to carboxylate free biotin was a unique property of the bacterial β-methylcrotonyl CoA carboxylase. However, Stoll et al. (140,311) demonstrated that the avian liver acetyl CoA carboxylase also catalyzes this model reaction, albeit at a very slow rate. Subsequently, the biotin carboxylase component of the E. coli acetyl CoA carboxylase system was found to carboxylate free biotin (96,100). Since this biotin carboxylase is free of the biotin-containing carboxyl carrier protein (100), it is evident that carboxylation occurs at a specific biotin site rather than by trans-carboxylation from the covalently linked biotin prosthetic group. This biotin site exhibits a low affinity for free biotin ($K_m = 170$ mM, biotin carboxylase (100); $K_m = 40$ mM, avian liver acetyl CoA carboxylase (140)) relative to that for the biotin-containing carboxyl carrier protein ($K_m \leq 2 \times 10^{-6}$ M) (97). In contrast, the carboxyl transferase component (E_b) of the E. coli acetyl CoA carboxylase system, which catalyzes the model carboxyl transfer reaction from malonyl CoA to free biotin, appears to have a much higher affinity ($K_m = 3.5$ mM) for free biotin (100a). The biotin-binding sites of all three carboxylases exhibit a high degree of stereochemical and structural specificity; however, there is a dissimilarity between the

* The report (143) that the 2'-carbonyl carbon atom of the biotinyl prosthetic group participates in carboxyl transfer was subsequently found to be incorrect (144).

bacterial specificity patterns and that of the avian enzyme (see Table III). The configuration about the 2-position, the point at which the side chain joins the tetrahydrothiophene ring, is critical since (−)-biotin is completely inactive. Furthermore, the importance of the tetrathiophene is indicated by the inactivity of dethiobiotin, biotin sulfoxides, and 2-imidazolidone, as well as the greatly reduced activity caused by substitution of O for S in the ring as with heterobiotin. Modification of the ureido group by replacing oxygen with sulfur as with 2'-thiobiotin completely abolishes activity. The length of the side chain is also a critical determinant of reactivity since the biotin side-chain homologues, norbiotin (C_4) and homobiotin (C_6), are much less reactive than biotin (C_5). Biocytin, which has a structure similar to the prosthetic group of the holoenzyme, is a poorer substrate than biotin for the bacterial carboxylases (81,100), although it is more active than biotin with the avian enzyme (140). This, and the fact that the maximal velocities for carboxylation of free biotin catalyzed by both bacterial carboxylases are about 1000-fold higher than with the animal carboxylase, indicate that the carboxylation sites of the bacterial and animal enzymes are significantly different. A possible role for sulfur in biotin function may be to facilitate binding of the bicyclic ring to the active site through hydrogen bonding. It has been demonstrated (312) that the sulfur atom is protonated by $FSO_3H \cdot SbF_5$, the "magic acid," trans with respect to the valeric acid side chain. Were an interaction with the enzyme to occur at this site, trans attachment would be predicted because of less steric hindrance by the imidazolidone ring and aliphatic side chain (see Fig. 4). Sigel et al. (313) have suggested that Mn(II) or Cu(II) interact with the tetra-hydrothiophene ring trans to the ureido ring and aliphatic side chain. The fact that O-heterobiotin (in which oxygen replaces sulfur in the tetrahydrothiophene ring) can be carboxylated (81,100,140), and can replace biotin for growth in some microorganisms (201,314–322), indicates that sulfur does not serve an essential catalytic function.

Biotin appears to participate in carboxylation reactions as a nucleo-philic catalyst. Model studies by Caplow and co-workers (118,323) using 2-imidazolidone as a biotin analogue revealed that the ureido N atom is a poor nucleophile in bimolecular reactions, raising funda-mental questions about its mechanism of action in carboxylation reactions. A low nucleophilicity was indicated by the failure of 2-imidazolidone to be acylated by p-nitrophenyl acetate, acetylimi-dazole, or acetyl-3-methylimidazolium chloride under a variety of

As pointed out earlier, it was established conclusively (26,81,82) that the methylated product of free biotin carboxylation is 1'-N-methoxycarbonyl biotin methyl ester. It has been suggested by Bruice and Hegarty (324) that were the initial site of carboxylation of biotin the ureido oxygen atom, following methylation, transfer of the methoxycarbonyl group to the 1'-N-position would be anticipated, giving rise to the more thermodynamically stable product. It is not apparent, however, why the sole methylated product of the enzymatic carboxylation of free biotin following this postulated nonenzymatic shift is the 1'-N-methoxycarbonyl biotin methyl ester rather than the 100 to 7 ratio of 1'-N- to 3'-N-methoxycarbonyl biotin methyl ester by chemical synthesis.

Another indication that carboxylation occurs at the 1'-N position of biotin rather than at the ureido oxygen is that chemically-synthesized 1'-N-carboxy-(+)-biotin and enzymatically-generated "free" carboxy-biotin exhibit similar decarboxylation kinetics (324a). The latter derivative, a product of transcarboxylation from malonyl CoA to free (+)-biotin catalyzed by the carboxyl transferase component of E. coli acetyl CoA carboxylase, undergoes pH-dependent decarboxylation at rates identical to those of authentic 1'-N-carboxy-(+)-biotin; at pH's 8.7, 7.8, and 6.1, the $t\frac{1}{2}$'s for decarboxylation at 22° are 133 min, 85 min, and 27 min, respectively (324a).

2. Decarboxylation and Carboxyl Transfer

Model studies with biotin analogues indicate that the reactivity of the N-carboxy group of enzyme-biotin-CO_2^- in decarboxylation and carboxyl transfer to nucleophilic acceptors is probably determined by electrophilic activation of the carboxylate group, nucleophilic activation of the acceptor molecule, and the basicity of the ureido system. It appears likely that all of these factors are involved in enzyme-mediated catalysis and that the transition states for enzymatic decarboxylation and transcarboxylation are probably similar. Chemical studies with model N-carboxy analogues of biotin have been carried out with particular attention paid to their tendency to decarboxylate. The stability of these N-carboxy derivatives relative to N-carboxy-biotin and enzyme-biotin-CO_2^- are summarized in Table IV. The decarboxylation rates of N-carboxyurea derivatives, such as N-carboxy-imidazolidone and N-carboxybiotin, are about an order of magnitude greater than the parent allophanate, indicating that the structure of

TABLE IV
Comparison of Decarboxylation Rates of Carboxybiotin Derivatives

Carboxybiotin derivative	Temperature	pH	$t\frac{1}{2}$ (min)	k_1 (min^{-1})	References
A. Free derivative					
N-Carboxyurea	~25°	8	1500	4.6×10^{-4}	325
N-Carboxyimidazolidone-2	~25°	8	175	4.0×10^{-3}	325
1'-N-Carboxybiotin	~25°	≥7.6	105	6.6×10^{-3}	325
B. Enzyme-biotin-CO_2^- derivative:					
Acetyl CoA carboxylase	~25°	7.5	20	0.034	139,309
+ Citrate (10 mM)	~25°	6.9	6.3	0.110	139,309
+ Acetyl CoA (0.5 μM)	~25°	7.5	~11	~0.063	309
+ Acetyl CoA (5 μM)	~25°	7.5	~1.8	~0.38	139
+ Acetyl CoA (0.5 μM)	~25°	7.5	~1.1	~0.63	309
+ Citrate + acetyl CoA (5 μM)	~25°	7.5	<0.3	<2.0	139
Propionyl CoA carboxylase	30°	7.5	14	0.05	83
+ ATP	25°	8.1	2.6	0.27	83
+ ATP + propionyl CoA	30°	7.5	~3	~0.23	83
β-Methylcrotonyl CoA carboxylase	26°	8.8	14	0.05	98
+ β-Methylcrotonyl CoA	26°	8.5	~4	~0.17	98
Transcarboxylase	30°	8.8	2.2	0.31	87
Methylmalonyl CoA decarboxylase	~25°	6.7	$\geq 6.9 \times 10^{-4}$ [a]	≥1000[a]	117

[a] Estimated assuming that the terminal step in the reaction is Enz–biotin–CO_2^- → Enz–biotin + CO_2; the reaction took place in the presence of methylmalonyl CoA, propionyl CoA being removed by the coupled assay system (116).

353

the prosthetic group is a factor determining reactivity. The decarboxyl-ation of the N-carboxy-2-imidazolidone anion has been shown by Caplow and Yager (118) to be a unimolecular process. The slow de-carboxylation of this anion relative to other carbamates (118) and its insensitivity to general or specific acid catalysis is apparently due to the poor leaving ability of the imidazolidone anion and the low basicity of the ureido system. These properties would guarantee the existence of enzyme-bound carboxybiotin in an aqueous environment. The stability of the N-carboxybiotinyl group to decarboxylation would appear to be advantageous during its translocation between catalytic sites (Section V) and would permit more selectivity in the subsequent carboxylation (323,326). Protonation of N-carboxybiotin or N-carboxyimidazolidone markedly increases its rate of decarboxylation via a process involving concerted proton transfer yielding the isourea form, as pointed out by Knappe (146):
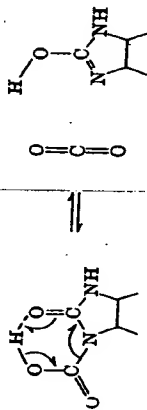
In view of the fact that the decarboxylation rate of carboxybiotin increases dramatically below pH 7.5 (146), it is conceivable that the enzyme could facilitate a similar protonation of carboxybiotin through general acid catalysis. It has also been demonstrated that divalent metal ions retard the decarboxylation of carboxyimidazolidone (118).

Incorporation of the N-carboxybiotin molecule into the active site of an enzyme leads to a further rate enhancement, ranging from 1 to 2 orders of magnitude, in the case of the abortive decarboxylation carried out by the carboxylases and transcarboxylase, to at least 6 orders of magnitude for methylmalonyl CoA decarboxylase (Table IV). Since decarboxylation is the *raison d'être* for the latter enzyme, it is not surprising that such a rapid turnover rate is obtained. Galivan and Allen (117) have investigated the temperature dependence of the overall reaction with methylmalonyl CoA decarboxylase, and have found an activation energy of 10.8 kcal. In contrast, Wood et al. (87) found with transcarboxylase that the activation energy for decarboxyla-tion of enzyme-biotin-CO_2^- is 26.6 kcal, the difference in activation energy between the two enzymes probably reflecting the predictably
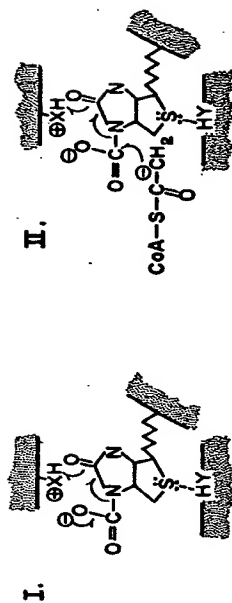
Fig. 9. Facilitation of enzyme-biotin-CO_2^- decarboxylation (I) and carboxyl transfer (II) by general acid catalysis.

greater stability of enzyme-biotin-CO_2^- for transcarboxylase. The binding of acceptor substrates to β-methylcrotonyl CoA, propionyl CoA, and acetyl CoA carboxylases completes the requirements for carboxyl transfer and enhances the decarboxylation of enzyme-biotin-CO_2^- at least 3- to 7-fold (Table IV). Furthermore, allosteric effectors of avian liver acetyl CoA carboxylase, such as citrate and isocitrate, increase the reactivity of the labile carboxyl group, and are synergistic with acetyl CoA, substrate and effector promoting a >20-fold increase in decarboxylation rate. It can be visualized (139) that activation resulting from conformational changes induced by the activator brings a proton-donating group into closer proximity with the carboxylated ureido system (Fig. 9), thereby facilitating general acid catalysis. The importance of the native conformation of the enzyme is indicated by the reduction in decarboxylation rate obtained by partial tryptic digestion with β-methylcrotonyl CoA carboxylase (326) or by dissociation of acetyl CoA carboxylase with 0.5 M NaCl (327).

Another factor which could be important in the electrophilic activa-tion of the N-carboxy group is the change in bond angle or length, resulting from strain or distortion induced by conformational changes at the active site. As pointed out by Jencks (328), "The induction of strain in the enzyme-substrate complex involves the overcoming of part of the energetic and entropic barrier to reaction by bringing the substrate part way along the reaction coordinate toward the transition state . . . ." In the case of carboxybiotin, electrophilic activation might be accomplished by deformation of the N-carboxy group, rendering it similar to a more electrophilic species, for example, CO_2, in the transition state. It has been demonstrated that the introduction of a single carboxyl group at the 1'-N position of the biotinyl moiety

induces sufficient conformational strain in liver acetyl CoA carboxylase so that the polymerized active form dissociates into protomers. The binding of its allosteric activator, citrate, prevents this dissociation presumably by constraining the enzyme in an active conformation compatible with the polymeric state. Under conditions where the enzyme is not carboxylated the K_D is 3 μM; however, under assay conditions, that is, where the enzyme is carboxylated, its activator constant, K_A, is 3 mM. This 1000-fold difference between K_D and K_A may be equivalent to that fraction of the binding energy utilized for the structural constraint needed to maintain the active conformation, notwithstanding its carboxylated state. Furthermore, it has been found by Edwards and Lane (329) that avidin (Section VII.A), which binds free- or enzyme-bound biotin with remarkable affinity ($K_{D\text{ free biotin}} = 10^{-15} M$) (5), enhances the decarboxylation of enzyme–biotin–CO$_2^-$. The $t\frac{1}{2}$ of enzyme–biotin–CO$_2^-$ (avian liver acetyl CoA carboxylase) at 2°, pH 7.5, of about 200 min is reduced to <1 min by binding to avidin. It is suggested that this effect may result from strain induced in the carboxyureido system owing to the extraordinarily tight binding of the carboxybiotinyl group by avidin.

By determining the equilibrium constant for the carboxylation of transcarboxylase by methylmalonyl CoA, and using approximations

$$\text{methylmalonyl CoA} + \text{Enz–biotin} \rightleftharpoons \text{propionyl CoA} + \text{Enz–biotin–CO}_2^-$$
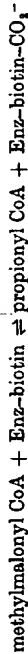
based on literature data, Wood et al. (87) calculated a $\Delta F^{0'}$ of -4.74 kcal/mole for the decarboxylation of enzyme–biotin–CO$_2^-$. This $\Delta F^{0'}$ is in the lower range for "energy rich" compounds.

The transcarboxylation reaction per se has not been studied successfully with model reactions. Schaeffer and Bhargava (330) have demonstrated that the model compounds N-arylcarbamyl-2-imidazolidone and N-alkyloxycarbonyl-2-imidazolidone transferred their carbonyl groups to nucleophiles such as piperidine and morpholine, although the reactivity of the model compounds was low. Studies by Knappe and Lynen (331) bearing on the ability of biotin to serve as a leaving group indicate that whereas alkaline hydrolysis of N-methoxycarbonyl-2-imidazolidone and methoxycarbonyl urea proceeds primarily via methoxide elimination, hydrolysis of N-methoxycarbonyl biotin proceeds to the same extent via both methoxide and biotinyl amide anion elimination. Thus it appears that biotin is a good leaving group which should enhance the reactivity of carboxybiotin in transcarboxylation reactions. The suggestion (155) that the sulfur atom of the

tetrahydrothiophene ring may participate in a transannular interaction with the ureido carbonyl group by donating electrons from its filled p-orbitals appears to have been ruled out (332,333). However, Bowen and Ingraham (334) suggest, on the basis of small chemical shifts (NMR) for the methylene protons adjacent to sulfur in solvents of varying dielectric constant and on the basis of other considerations, that the sulfur atom may in fact accept electrons from the carbonyl group into its empty d-orbitals. This effect might be promoted by hydrogen bonding or protonation of the ring sulfur atom by the enzyme (312, 334,335); hence enzymatic control could be exerted over the reactivity of the ureido system by altering the hydrogen bonding of the sulfur (334).

C. CARBOXYL ACCEPTORS

1. Acyl CoA Derivatives

Two types of carboxyl acceptors—acyl CoA derivatives and α-keto acids—have been investigated in some detail. The biotin-dependent enzymes having keto acid substrates utilize tightly bound manganese (157), zinc (122,336), and cobalt (122) for substrate activation. Acyl CoA carboxylases, which do not appear to contain tightly bound metal ions (337,338), must rely on other enzyme substituents for activation of their thioester substrates. Investigations with these carboxylases show that proton release from the α-position, as measured by ³H exchange using ³H-labeled water or substrate, depends upon the presence of the carboxylated biotinyl prosthetic group (339–342). Furthermore, the reaction catalyzed by propionyl CoA carboxylase occurs with retention of configuration, that is, carboxylation from the same side as proton abstraction (339,343–345). These observations and the fact that the reaction proceeds without a primary ²H or ³H kinetic isotope effect, when (R)- or (S)-2-deuterio- or 2-tritio-propionyl CoA is used as substrate (339,346), are consistent with a concerted mechanism in which carboxylation and proton removal occur simultaneously as shown below. A similar mechanism has been proposed
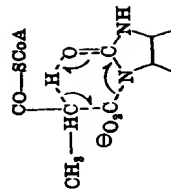
(155) for the metal-dependent biotin enzymes (Section IV.C.2). The above results are inconsistent with a mechanism for acyl CoA carboxylation in which carbanion formation is rate-determining and subsequent proton release from the enzyme and carboxylation are rapid. However, other kinetic variations of a carbanion mechanism are possible.

Interesting in this regard are the recent observations that several biotin enzymes [liver acetyl CoA carboxylase, propionyl CoA carboxylase, and the biotin-free carboxyl transferase component of E. coli acetyl CoA carboxylase) catalyze a slow avidin-insensitive decarboxylation of malonyl CoA derivatives (347,100a). The occurrence of this abortive reaction indicates that labilization of the β-carboxyl group of the malonyl CoA derivative and insertion of a proton in its place need not involve biotin. However, in the presence of free biotin E. coli carboxyl transferase catalyzes transcarboxylation from malonyl CoA forming free N-carboxybiotin (100a).

Convincing evidence for a carbanion (or enolate) mechanism involving acyl CoA substrates has been provided by Eggerer (348,349) with citrate synthase (E.C. 4.1.3.7) and malate synthase (E.C. 4.1.3.2). Earlier investigations (350,351) failed to demonstrate significant synthase-catalyzed deuterium or tritium exchange between the methyl hydrogens of acetyl CoA and water. Eggerer (348) noted that tritium exchange from acetyl CoA was facilitated by (S)-malate, presumably by mimicking the effect of oxaloacetate, the condensing partner in the reaction. However, (R)-malate was inactive, indicating a high degree of stereospecificity for the mediator analogue. This effect, referred to as "substrate synergism" by Bridger et al. (352), is probably brought about by (S)-malate-induced conformational changes at the active site which promote catalysis or binding of the second substrate, acetyl CoA.

Model studies indicate that the thioester carbonyl increases the acidity of the α-hydrogen atom. Lynen (353) has shown that the α-hydrogens of β-ketothioesters are more labile than those of β-ketoesters. The pK_a for acetoacetate is 12.7, that for ethylacetoacetate 10.7, and that for (S)-acetoacetyl-N-acetoacetyl-2-thioethylamine is 8.5. Activation of the α-hydrogen of thioesters is presumed to arise from their ability to resonance stabilize the resulting carbanion. In comparison to the carbonyl group of oxygen esters, that of thioesters has greater double-bond character because of the inability of sulfur to participate in electron pair-donor type resonance with the same

facility as oxygen (see below) (350,354,355). Baker and Shulgrin

$$\underset{\mathbf{1}}{R-\overset{\overset{\displaystyle O}{\|}}{C}-O-R} \leftrightarrow \underset{\mathbf{2}}{R-\overset{\overset{\displaystyle O^-}{|}}{C}=\overset{+}{O}-R}$$
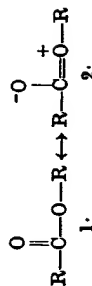
(356) have shown that the carbonyl oxygen atom in thioesters carries less negative charge than that in the corresponding esters, suggesting that resonance structure 2 is less important in thioesters. These studies indicate that the thioester carbonyl function weakens the adjacent carbon–hydrogen bond, thereby rendering acyl CoA derivatives more susceptible to electrophilic attack.

2. α-Keto Acids

All biotin-dependent α-keto acid carboxylases (Section VI.E), transcarboxylases (Section VI.G), and β-keto acid decarboxylases (Section VI.I) appear to be metalloenzymes. Mildvan, Scrutton, and Utter (155) have demonstrated that tightly bound manganese functions in the activation of pyruvate by pyruvate carboxylase, whereas transcarboxylase studied in Wood's laboratory (122) utilizes either cobalt or zinc for α-keto acid activation. Recently, oxaloacetate decarboxylase partially purified by Stern (115) was shown to be inhibited by metal binding reagents such as cyanide and 8-hydroxyquinoline, suggesting that it too is a metalloenzyme. The existence of an enzyme–manganese–pyruvate bridge complex has been unequivocally established with pyruvate carboxylase, pyruvate acting as a monodentate ligand of the bound metal (Fig. 10) (337). It was observed that enzyme-bound manganese markedly broadened the nuclear magnetic resonance line of the methyl protons of pyruvate (337). Assessment of the paramagnetic contribution to the relaxation rate of the methyl protons of pyruvate revealed a marked enhancement owing to their interaction with the unpaired electrons of enzyme-bound manganese. This strongly suggests that the metal promotes a weakening of the α-hydrogens of pyruvate, thereby increasing the nucleophilic character of the methyl group. However, no evidence for carbanion formation has been obtained. Tritiated pyruvate only exchanged with water protons in the presence of the complete reaction mixture or oxaloacetate; hence proton activation occurs only when the biotin prosthetic group is carboxylated (155). Carbanion formation cannot be ruled out, since the proton released from pyruvate may be

retained by the enzyme and reinserted into pyruvate without exchange with water protons. Precedent for proton retention of this sort has been established with several enzymes (357,358).

Further evidence for the participation of tightly bound metal ion in the enzymatic activation of keto acids has been obtained with trans-carboxylase by Northrop and Wood (122). A similar broadening of the NMR signal of the methyl protons of pyruvate by the metalloenzyme was observed. In both cases oxalate, an uncompetitive inhibitor with respect to pyruvate and a competitive inhibitor with respect to oxaloacetate, reverses the broadening of the NMR signal (122,337). A concerted mechanism has been proposed by Mildvan et al. (Fig. 10) (155) for the manganese-mediated carboxyl transfer from enzyme–biotin–CO_2^- to pyruvate catalyzed by pyruvate carboxylase. A similar mechanism for the cobalt-mediated carboxyl transfer catalyzed by transcarboxylase has been suggested by Northrop and Wood (122). The former mechanism visualizes the synchronous abstraction of a methyl proton coupled with carboxyl transfer from the 1'-N-position of biotin. Although there is convincing evidence for the nucleophilic activation of pyruvate by the tightly bound metal, evidence is lacking for a role of the metal in the electrophilic activation of the N-carboxy group. The recent findings of Rose (359) with pyruvate carboxylase show that, like the reaction with propionyl CoA carboxylase (339,343, 344), carboxylation proceeds with retention of configuration. Further-more, the pyruvate carboxylase-catalyzed reaction was found (359) to exhibit a primary kinetic isotope effect discriminating against ^3H and ^2H in the C-3 position of pyruvate. This suggests that proton removal may occur in the rate-limiting step of the reaction, which is consistent with the concerted mechanism shown in Figure 10, or that the proton removed does not freely exchange with water protons.

V. Kinetic and Molecular Basis for Prosthetic Group Translocation

A. KINETICS: "PING-PONG"

The isolation of an enzyme–biotin–CO_2^- intermediate which participates in both half-reactions (Section IV.B.1; Fig. 2), inhibition of both partial reactions by the biotin-binding protein avidin (Section II.B.2), and the inability of apoenzymes to catalyze the overall reaction (Section III.C.) (254,255,260) have led to a mechanism in which carboxyl transfer occurs via the biotin prosthetic group. The
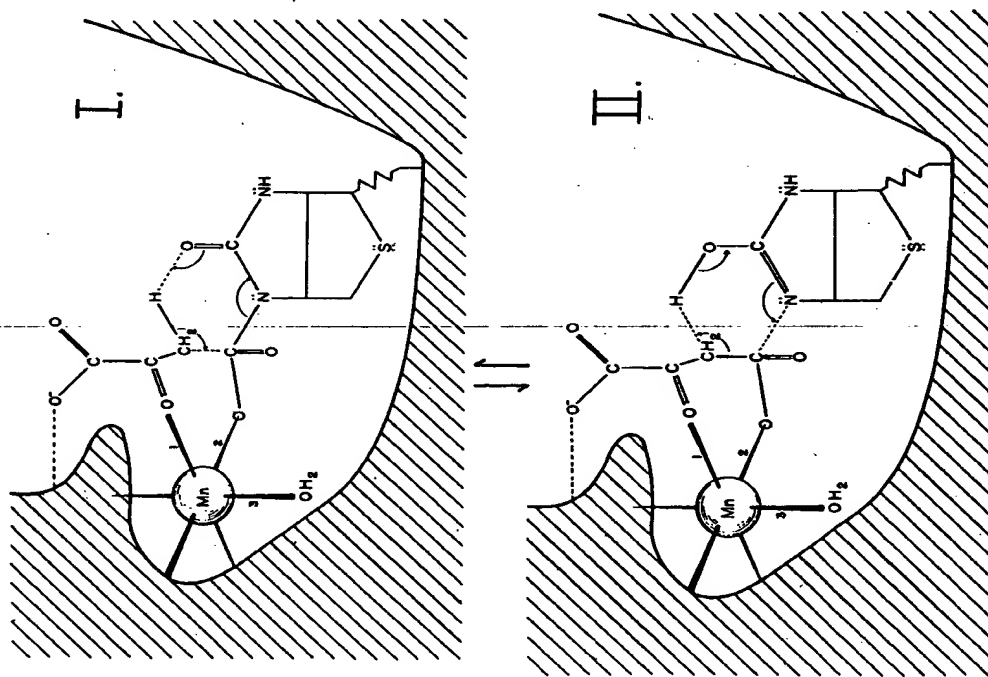
Fig. 10. Proposed role for enzyme–manganese–substrate bridge complexes in the carboxylation of pyruvate (I) and decarboxylation of oxaloacetate (II).

biotinyl moiety is assumed to load one carboxyl group at an initial site (Fig. 11, site I), flip to a neighboring site (Fig. 11, site II), and then transfer its "activated CO_2" to an appropriate acceptor. This model, which corresponds to the classical "ping-pong" mechanism, has been shown on first approximation to apply to the transcarboxylase- and acetyl CoA carboxylase-catalyzed reactions (90–93) and probably holds for other biotin enzymes. However, Northrop (90) has demonstrated with transcarboxylase that ternary complex formation (involving substrates of both partial reactions) also occurs, suggesting that a hybrid "ping-pong" mechanism is a more appropriate description of the kinetic events. A classical "ping-pong" mechanism (360–362) would involve release of the products (C and D, mechanism 15) of the loading reaction prior to the binding of the substrate (E, mechanism 15) of the partial reaction. In the case of bacterial transcarboxylase (90,

$$\text{Enz–biotin} \quad \overset{A\ B}{\downarrow\downarrow} \quad \overset{C\ D}{\uparrow\uparrow} \quad \text{Enz–biotin–CO}_2^- \quad \overset{E\ F}{\downarrow\downarrow}\uparrow \quad \text{Enz–biotin} \qquad (15)$$
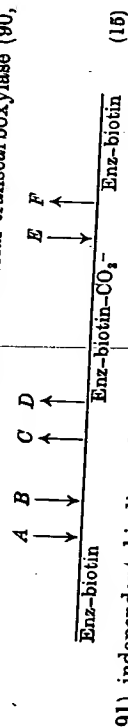
91) independent binding of substrates of both partial reactions can occur to some extent, that is, substrates of the second half-reaction can bind prior to the release of the products of the first half-reaction, thus making possible, but not necessitating, ternary complex formation.
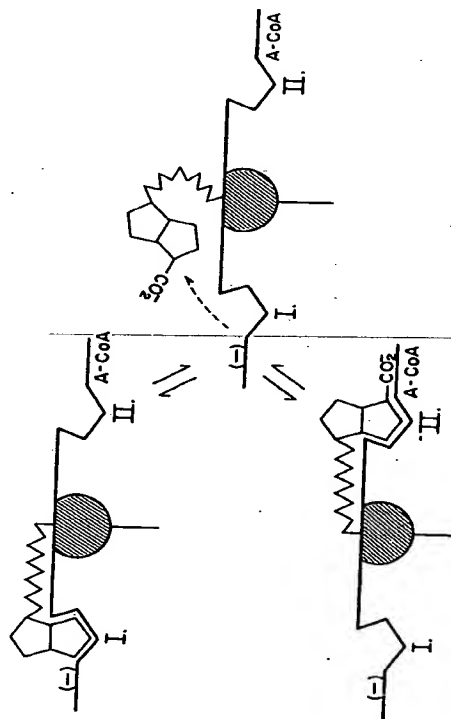
Fig. 11. Translocation of the carboxylated biotinyl prosthetic group.

Nevertheless the two substrate binding sites are not totally independent, since substrate binding is influenced by the state of carboxylation of the biotin prosthetic group. The carboxybiotinyl group would thus compete in essence with carboxylated acceptor for a common site (shown below), but allow the binding of uncarboxylated acceptors.
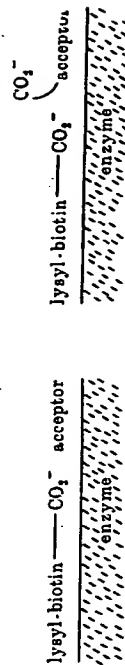
In the case of rat-liver acetyl CoA carboxylase, Numa and co-workers (92,93) have obtained kinetic evidence which indicates that carboxylation of biotin proceeds by an ordered mechanism with ATP binding prior to HCO_3^-, and following carboxybiotin formation, P_1 release subsequent to ADP. The reaction is postulated to proceed via a "bi bi uni uni ping-pong" mechanism as shown in mechanism 16. It seems

$$\text{Enz–biotin} \quad \overset{\text{ATP}\ \ HCO_3^-}{\downarrow\ \ \downarrow} \quad \overset{\text{ADP}\ \ P_1}{\uparrow\ \ \uparrow} \quad \text{Enz–biotin–CO}_2^- \quad \overset{\text{acetyl}\ \ \text{malonyl}}{\underset{\text{CoA}\ \ \ \text{CoA}}{\downarrow\ \ \uparrow}} \quad \text{Enz–biotin} \qquad (16)$$
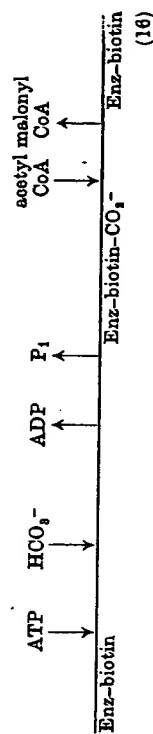
likely that other reactions catalyzed by biotin-dependent enzymes would also proceed via "ping-pong" mechanisms.

B. MOLECULAR BASIS

The "ping-pong" kinetic mechanism implicates separate sites for the substrates of both half-reactions, as well as for the mobile biotinyl group. Recent work on acetyl CoA carboxylase from *E. coli* (Section VI.A) and transcarboxylase from *P. shermanii* (Section VI.G) indicates that these sites may reside on different subunits. Alberts et al. (96) have resolved the acetyl CoA carboxylase system into three functional components: (a) a carboxyl carrier protein (CCP) which contains the covalently-bound biotin prosthetic group, (b) biotin carboxylase, which catalyzes the ATP-dependent carboxylation of CCP-bound or free biotin, and (c) an additional protein, E_b, which was presumed to have a role in carboxyl transfer from the carboxylated prosthetic group to acetyl CoA. The role of E_b was recently defined in Lane's laboratory (100a) where it was recently shown that this component catalyzes

carboxyl transfer from malonyl CoA to free d-biotin to form acetyl CoA and N-carboxybiotin. As visualized in Figure 11, the biotin-free BC component (I in Fig. 11) catalyzes the carboxylation of the biotinyl prosthetic group on CCP (cross-hatched area in Fig. 11). Following the translocation of the carboxylated biotinyl group from site I to site II, carboxyl transfer to acetyl CoA is thought to be catalyzed by CT (II in Fig. 11). In Wood's laboratory a similar biotin-containing carboxyl carrier protein has been isolated from transcarboxylase (121). This enzyme has been resolved into several catalytically inactive components, one of which is a CCP (121,123). The function of the other components has not yet been determined, but all of these are required for reconstitution of transcarboxylase activity. Consistent with the occurrence of a distinct biotin-containing CCP in the bacterial enzymes is the fact that the avian liver acetyl CoA carboxylase, which is composed of four subunits, contains only one covalently bound biotin (119,311). However, in the case of the avian liver enzyme, the molecular weight of the biotin-containing carboxyl carrier protein is 100,000, rather than 9000 or 12,000 as with the $E.$ $coli$ and $P.$ $shermanii$ CCPs, respectively. Further investigations will be necessary to ascertain whether the apoenzyme components, for example, BC and CT, are specific for their homologous carboxyl carrier proteins.

The fact that the catalytic sites appear to reside on different subunits, distinct from the carboxyl carrier protein, indicates that either the prosthetic group must oscillate between sites or the subunits bearing these sites must move with respect to the prosthetic group. Since the functional bicyclic ring resides at the distal end of a flexible 14-Å side chain which anchors it to the apo CCP, the biotinyl group is capable of migrating between remote sites (91,119,155).

VI. Properties of the Biotin-Dependent Enzymes

A comparison of the molecular and catalytic properties of the biotin-dependent enzymes is found in Table VII, Section VI.J. For further information regarding comparative aspects and general properties of these enzymes, Sections II-V should be consulted.

A. ACETYL CoA CARBOXYLASE

1. Background

Acetyl CoA carboxylase was first recognized as one of two essential enzymatic components (R_1 and R_2) for fatty acid synthesis from

acetyl CoA catalyzed by avian liver extracts (67-69,102,103,363-375). The other component, fatty acid synthetase, utilizes malonyl CoA generated by the carboxylase for condensation with the fatty acyl CoA being elongated (101,376,377). The observation that fatty acid synthesis required bicarbonate (67,368,369,378) led to the discovery (102) that R_1, one of the essential enzymatic components of the system, catalyzes an ATP- and acetyl CoA-dependent carboxylation reaction. It was suggested by Lynen (379) and later demonstrated by Wakil (102) and Formica and Brady (103) that the enzymatic carboxylation product was malonyl CoA. Acetyl CoA carboxylase was identified as a biotin-dependent enzyme by Wakil and co-workers (67-69,371), who observed that the biotin content of carboxylase preparations increased during purification (67,68) and that catalytic activity could be blocked by avidin, but not by biotin-saturated avidin (68,69). The "carboxyl-carrier" role of the biotin-prosthetic group of this enzyme has since been unequivocally established (88,89,95,96,100,119,136,144,154,311).

In 1952 Brady and Gurin (380) demonstrated that the synthesis of fatty acids from acetate in pigeon-liver extracts was markedly activated by tricarboxylic acids It became evident that cell-free animal systems, including those from avian liver, rat liver, adipose tissue, and mammary gland, required tricarboxylic acid activators, notably citrate and isocitrate, for maximal rates of fatty acid synthesis from acetate or acetyl CoA (363-366,381-390). Once the enzymatic components of the synthetic system had been resolved and the chemical steps elucidated (101-103,368,369,372-375), investigators in several laboratories independently demonstrated (131,388,392-398) that the site of citrate activation is the acetyl CoA carboxylase-catalyzed reaction. Vagelos et al. (399,400) made the interesting observation that activation of the carboxylase was accompanied by its increased rate of sedimentation on sucrose density gradients. This increased sedimentation velocity was subsequently shown by Gregolin et al. (119,401,402) to be due to the polymerization of the protomeric form of the carboxylase which gives rise to unique filamentous structures. The activation phenomenon has since been investigated extensively with purified acetyl CoA carboxylases from many animal species and diverse organ systems (401-409). Activation by citrate and concomitant polymerization to give filamentous structures appears to be characteristic of the acetyl CoA carboxylases of animal origin (406). Although the carboxylase from aerobically

grown yeast is activated by citrate, it does not undergo polymerization (410–412). Furthermore, the carboxylases from plant sources (124,413, 414), *E. coli* (95,100), and brewer's yeast (131,154) are insensitive to tricarboxylic acid activators.

2. Partial Reactions

The reaction catalyzed by acetyl CoA carboxylase involves the minimal 2-step reaction sequence shown below. The participation of

$$\text{Enz-biotin} + \text{HO—CO}_2^- + \text{ATP} \underset{\text{Mg}^{2+} \text{ or Mn}^{2+}}{\rightleftharpoons} \text{Enz-biotin-CO}_2^- + \text{ADP} + P_i \quad (17a)$$

$$\text{Enz-biotin-CO}_2^- + \text{acetyl CoA} \rightleftharpoons \text{Enz-biotin} + \text{malonyl CoA} \quad (17b)$$

Net:

$$\text{ATP} + \text{HO—CO}_2^- + \text{acetyl CoA} \underset{\text{Mg}^{2+} \text{ or Mn}^{2+}}{\rightleftharpoons} \text{malonyl CoA} + \text{ADP} + P_i \quad (17c)$$
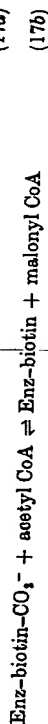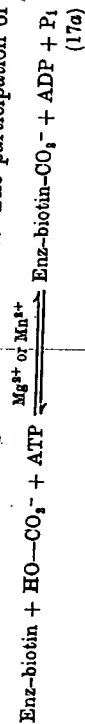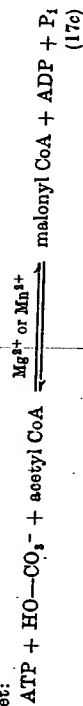
the first step (reaction 17a) in the overall carboxylation reaction is supported by carboxylase-catalyzed P_i, HCO_3^-, and divalent cation-dependent ATP-^{14}C-ADP exchange (415). The reciprocal ATP^{32}-P_i exchange also catalyzed by the enzyme requires ADP, HCO_3^-, and divalent cation (125,131,136). Precedent for the occurrence of both of these exchange reactions had been established previously with β-methylcrotonyl CoA (81) and propionyl CoA (129) carboxylases. Furthermore, as with these enzymes (83,86), the carboxylated form (Enz-biotin-CO$_2^-$) was isolated following incubation with ATP, HCO_3^-, and divalent metal ion (97,119,139) or with malonyl-3-^{14}C-CoA (88). As with other biotin-dependent enzymes (84–87), the site of carboxylation was found to be the 1'-N position of the biotinyl prosthetic group (Section IV.B.1) (88,89,144). The avian liver (140,311) and *E. coli* (96,100) carboxylases have also been shown to catalyze the ATP- and divalent metal ion-dependent carboxylation of free (+)-biotin, a model reaction (80) for the first step (reaction 17a). The carboxylated product of this reaction was identified as 1'-N-carboxybiotin (96,100, 140). Support for the participation of an enzyme-biotin-CO$_2^-$ intermediate in the overall reaction, as well as for the second partial reaction, is the ability of the isolated carboxylated enzyme to transfer its carboxyl group to acetyl CoA or acetyl pantetheine (88,97,139,140). Furthermore, transcarboxylation from malonyl CoA to free (+)-biotin forming free (+)-biotin-CO$_2^-$, a model for the second partial reaction, is

catalyzed by the carboxyl transferase component of *E. coli* acetyl CoA carboxylase (100a). In addition, malonyl CoA-^{14}C-acetyl CoA exchange, consistent with the occurrence of the second partial reaction, has been obtained with the avian (136,402), yeast (154), wheat-germ (151), and rat-liver (131) enzymes. All of the preceding reactions are inhibited by avidin (125,131,136,151,154) and, in the case of the animal enzymes, markedly activated by tricarboxylic acids, notably citrate and isocitrate (131,136).

More recently, Numa and co-workers (92,93) have shown that the kinetic patterns, obtained in initial velocity and product inhibition studies with the rat-liver enzyme, are consistent with the 2-step reaction sequence shown above (reactions 17a, b). An ordered sequence is visualized for the first partial reaction in which ATP binds initially, followed by HCO_3^-; subsequent to carboxybiotin formation, ADP is released, followed by P_i. It has been proposed (92,93) that the carboxylation reaction occurs via a "bi bi uni uni ping-pong" mechanism (Section V.A).

3. Molecular Characteristics

Acetyl CoA carboxylases purified from several animal tissues exist as enzymatically active polymeric filaments of high molecular weight (401,406,416–418) and have similar electron microscopic, hydrodynamic, and catalytic properties (406). Electron microscopic examination of homogeneous preparations of avian liver and bovine perirenal adipose tissue (Fig. 12) reveal that their filamentous structures are virtually indistinguishable (406). Similar filamentous forms of the rat- (419) and human-liver (406) carboxylases have also been observed. This high degree of structural organization exhibited by the animal carboxylases suggests a possible structural role in addition to their known catalytic and regulatory functions. Conceivably, the carboxylase filaments could serve as an organizing matrix for a loose supramolecular complex of other enzymes taking part in lipid biosynthesis. Cytofilaments with dimensions similar to those of the carboxylase filaments have been observed surrounding the triglyceride droplets in thin sections of developing and mature adipose tissue cells (420–422). Whether these intracellular filaments are identical to the acetyl CoA carboxylase filaments isolated from adipose tissue is a matter of conjecture.
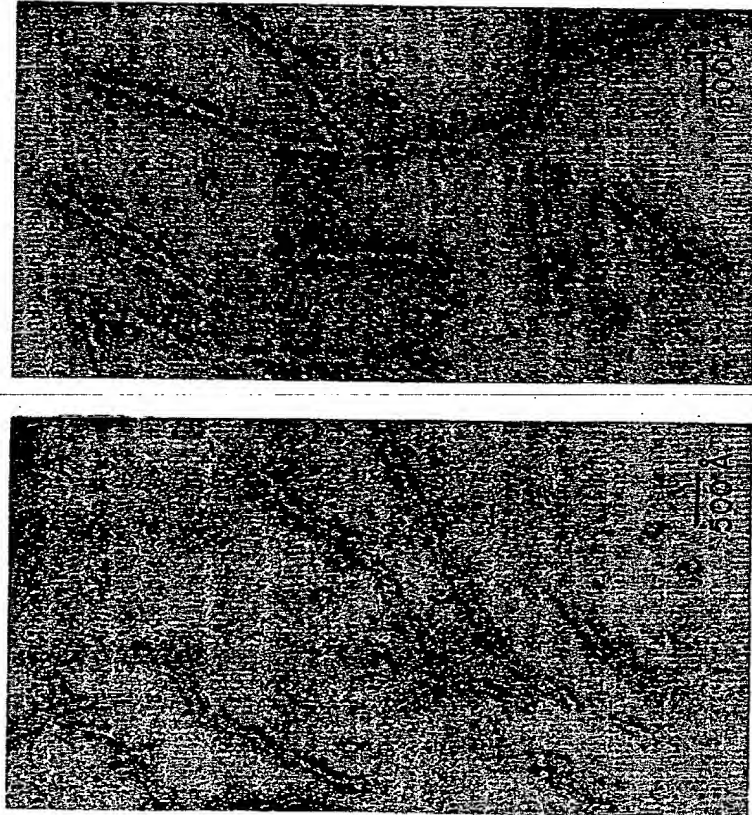
Fig. 12. Filamentous forms of acetyl CoA carboxylases (a) from avian liver and (b) from bovine adipose tissue, in the presence of citrate. Dilute solutions (20 μg/ml) of chicken liver or bovine perirenal adipose tissue carboxylase in 50 mM tris (Cl⁻) buffer contained 10 mM potassium citrate, 5mM 2-mercaptoethanol, and 0.1 mM ethylenediaminetetraacetate at pH 7.5. The preparations were stained with 4% aqueous uranyl acetate. From Kleinschmidt et al. (406).

The filaments from avian liver and bovine adipose tissue have a twisted appearance with indentations along the longitudinal axis suggestive of a helical structure (Fig. 12) (406). The widths of the filaments ranged from 70 to 100 Å with lengths up to 5000 Å. A 130-140 Å period is visible both in individual filaments (406) and in paracrystalline fibers (418) of the avian enzyme. A filamentous structure is consistent with their hydrodynamic properties, including the high intrinsic viscosity ($[\eta] = 100$, avian liver carboxylase) (98),

the hypersharp character of the sedimenting boundary in the analytical ultracentrifuge (401), and the marked concentration dependence of $s_{20,w}$ and η_{sp}/c (98,415). The carboxylase filaments are linear polymers, having molecular weights of several million (Table V), and are composed of weight-homogeneous protomers (406). The protomeric species have weights of 410,000 daltons ($s_{20,w} = 13.1$ S) (402,423-425) in the case of the avian liver enzyme and approximately 550,000 daltons ($s^{\circ}_{20,w} = 14.7$ S) for the bovine adipose tissue enzyme (416). These protomers also appear to be asymmetric structures as shown by viscosity studies and by the high dependence of $s_{20,w}$ on enzyme concentration (347,416). Dissociation of the avian liver carboxylase with sodium dodecyl sulfate gives rise to weight-homogeneous subunits of about 100,000 daltons (119,311,423), which indicates that the protomeric unit is composed of 4 subunits. Since the enzyme contains 1 mole of covalently bound biotin, has one carboxylation site, and binds tightly 1 mole of citrate and 1 mole of acetyl CoA per 410,000 g of protein (119,311,402, 424), it is evident that the subunits are nonidentical. Each protomer contains a single biotin prosthetic group, which indicates that they are identical (311). As will be indicated later (Section VI.A.4), carboxylase activity is determined by the state of the protomer-polymer equilibrium, catalytic activity being associated only with the polymeric species.

In contrast to the acetyl CoA carboxylases from animal tissues (131,136,397,400,407,408) and yeast (154,410-412) which can be obtained as stable multiple subunit structures, the E. coli carboxylase dissociates readily into three functionally distinct protein components (95,96,100,100a,100b) all of which are essential for the carboxylation of acetyl CoA (95-97,100a). These components include (a) biotin carboxylase (BC) which is free of biotin and catalyzes the carboxylation of the carboxyl carrier protein or free (+)-biotin (96,97,100), (b) a biotin-containing carboxyl carrier protein (CCP) (97), and (c) a protein, "carboxyl transferase" (CT), presumed to catalyze carboxyl transfer to acetyl CoA (95). Biotin carboxylase has a molecular weight of approximately 100,000 ($s_{20,w} = 5.7$ S) and is composed of two 50,000-dalton subunits (100). The E. coli carboxyl carrier protein has a molecular weight of about 9,000 ($s_{20,w} = 1.3$ S) (97) which is similar to CCP (12,000 daltons, $s_{20,w} = 1.3$ S, 1 biotin/molecule) isolated from P. shermanii (Section VI.G.3) (121). The biotin carboxylase and carboxyl

carrier protein from *E. coli* have recently been crystallized in Lane's (100,100b,428) and Vagelos' (427) laboratories, respectively. The carboxylation of acetyl CoA catalyzed by the *E. coli* carboxylase system appears to proceed via the following reaction sequence:

$$\text{CCP-biotin} + \text{HO—CO}_2^- + \text{ATP} \underset{\text{BC}}{\overset{\text{Mg}^{2+}\text{ or Mn}^{2+}}{\rightleftharpoons}} \text{CCP-biotin-CO}_2^- + \text{ADP} + \text{P} \quad (18a)$$

$$\text{CCP-biotin-CO}_2^- + \text{acetyl CoA} \underset{\text{CT}}{\rightleftharpoons} \text{CCP-biotin} + \text{malonyl CoA} \quad (18b)$$

Net:

$$\text{ATP} + \text{HO—CO}_2^- + \text{acetyl CoA} \underset{\text{BC, CCP, CT}}{\overset{\text{Mg}^{2+}\text{ or Mn}^{2+}}{\rightleftharpoons}} \text{malonyl CoA} + \text{ADP} + \text{P}_i \quad (18c)$$
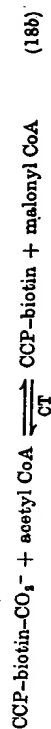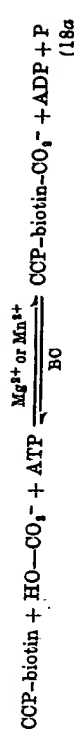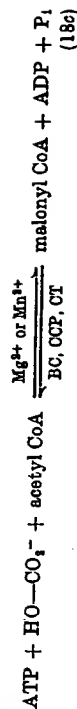
It has been demonstrated in Vagelos' laboratory (97) that in the presence of ATP, $H^{14}CO_3^-$, and divalent cation, biotin carboxylase catalyzes the carboxylation of CCP to form $\text{CCP-}^{14}\text{CO}_3^-$. Carboxyl transfer from the $\text{CCP-}^{14}\text{CO}_3^-$ preparation (still containing BC) to acetyl CoA forming malonyl CoA could be effected in the presence of CT. It was demonstrated by Dimroth et al. (100,100b) that the carboxylation of free (+)-biotin catalyzed by crystalline BC is activated 11-fold (V_{max} effect) by 15% (v/v) ethanol and certain other organic solvents. More recently Guchhait et al. (100a) have shown that CT catalyzes carboxyl transfer from malonyl CoA to free (+)-biotin forming free 1'-N-carboxy-(+)-biotin. Although neither enzymatic component, i.e., BC or CT, contains covalently-bound biotin, both catalyze model reactions with free biotin, which account for their respective catalytic functions in the half-reactions (Reactions 17a and 17b) of the overall process. Therefore, both the BC and CT components must contain distinct biotin binding sites. It is apparent that the biocyclio ring of the biotinyl prosthetic group of the carboxyl carrier protein oscillates between two catalytic sites, i.e., biotin carboxylating and carboxyl transferring, which are housed on different subunits (see Fig. 11).

The acetyl CoA carboxylase from wheat germ also dissociates during purification into what appears to be nonidentical components (124,125). However, further purification of the dissociated components, as well as reconstitution of carboxylase activity, will be necessary before functional roles can be assigned.

TABLE V

Comparison of the molecular and catalytic properties of acetyl CoA carboxylases isolated from bovine perirenal adipose tissue and chicken liver. The data were obtained from investigations in our laboratory (119,136,139,311,401,402,406).

Condition	Microscopic form	Carboxylation rate ($\mu M/min/mg$)	Sedimentation velocity		Molecular weight
			Sucrose density gradient (S)	Analytical ultracentrifuge (S)	
Bovine adipose tissue carboxylase					
Assay mix without citrate	Protomeric	≤0.16	13-15	—	—
0.5 *M* NaCl, pH 8	—	—	13-16	14.7	~550,000
Assay mix + citrate	Filamentous	12	47-50	—	
10 mM Citrate, pH 7.5	Filamentous	—	50-55	68	several million
Avian liver carboxylase					
Assay mix without citrate	Protomeric	≤0.30	13-15	—	—
0.5 *M* NaCl, pH 8	Protomeric	11	13-15	13.1	410,000
Assay mix + citrate	Filamentous	—	47-50	—	
10 mM Citrate, pH 7.5	Filamentous	—	50-55	69	4-5 million
0.1-1.0 % sodium dodecyl sulfate (subunit)		0	—	4	110,000

Biotin prosthetic group 0.93
Acetyl CoA binding sites 1.0 } per protomeric unit (410,000 daltons)
Citrate binding sites 1.16

4. Regulation

The reaction catalyzed by acetyl CoA carboxylase appears to be the initial committed step in the biosynthesis of fatty acids and other acetogenins (390,428–430). A large body of evidence supports the role of the carboxylase in the regulation of fatty acid synthesis in animal tissues (390), whereas little is known regarding its regulatory role in yeast, bacteria, and plant tissues. In animal systems the reaction catalyzed by the carboxylase occurs early in the extramitochondrial pathway of fatty acid synthesis (Fig. 13) and constitutes the first committed step, malonyl CoA having no other apparent metabolic alternative. This step has particular regulatory potential since it is effectively the rate-determining step of the synthetic process in animal tissues (377,388,390,431,432). In the absence of citrate, its tricarboxylic acid activator, acetyl CoA carboxylase activity is lower by nearly two orders of magnitude, than in the fully activated state where its catalytic capacity is nearly kinetically matched to those of the citrate cleavage enzyme and fatty acid synthetase (see Fig. 13) (390,431,433–435). Hence changes in either effector level or fluctuations in the tissue concentration of the carboxylase could control the rate of acetyl CoA carboxylation and, thereby, regulate fatty acid synthesis. The hepatic concentration of citrate has been reported (436–440) to vary in concert with changes in the rate of lipogenesis in different physiological states, although these findings have not been consistently obtained (441–445). Fatty acyl-CoA derivatives are end products of this biosynthetic pathway; hence a reasonable metabolic role can be visualized for these compounds as feedback inhibitors of an early step
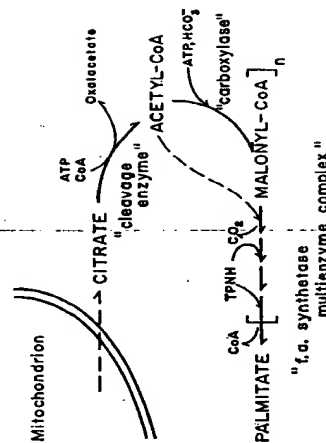
Fig. 13. Extramitochondrial pathway for fatty acid synthesis in animal tissues.

in the process. Such a role would be consistent with the observation that fasting (432,445–450), fat feeding (450–452), and alloxan diabetes (450,453–457) which are known to depress hepatic fatty acid synthesis, lead to an increased long-chain fatty acyl-CoA concentration in liver (449,450,458–463).

Several hypolipidemic drugs (2-methyl-2-phenoxypropionate derivatives) have been found to inhibit acetyl CoA carboxylase (464,465). All appear to inhibit competitively with respect to acetyl CoA and tricarboxylic acid activator and noncompetitively with respect to ATP and HCO_3^-.

Fluctuations in the level of acetyl CoA carboxylase in tissues are correlated with alterations in the rate of fatty acid synthesis in different nutritional and physiological states (390,431,451,457,466). The rise in carboxylase activity following refeeding of a high carbohydrate diet to fasted animals, which leads to increased lipogenesis (467) or the insulin-stimulated rise in lipogenesis, is blocked by inhibitors of protein synthesis, such as actinomycin D and puromycin (468–472). Majerus and Kilburn (473) and Numa et al. (93) have shown that synthesis and degradation of the carboxylase are responsive to physiological changes which alter the rate of fatty acid synthesis. However, different factors appear to govern the rates of synthesis and degradation of the enzyme (93,473). It is not known by what mechanism the carboxylase is degraded in vivo, although it is known that in vitro the catalytically inactive protomeric form of the enzyme is intrinsically much less stable and undergoes irreversible inactivation more rapidly than the active polymeric form (139,416).

a. Factors Controlling Protomer-Polymer Equilibrium. Acetyl CoA carboxylase from animal tissues appears to be capable of oscillating between catalytically active polymeric and catalytically inactive protomeric states, the level of carboxylase activity presumably governed by the position of the protomer-polymer equilibrium (Fig. 14) (119,399–403,406,474–476). Tricarboxylic acid activators, such as citrate and isocitrate, are capable of shifting this equilibrium toward the catalytically active state, apparently by inducing a productive conformational change in the carboxylase protomer which favors polymerization. This is indicated by the fact that under conditions approximating those of the enzymatic assay, citrate or isocitrate induce a change in $s_{20,w}$ from 13–15 S to 47–50 S (119,401,402,406), a rapid transition from protomeric to filamentous form as judged by electron

indicates that catalytic activity is determined by the polymeric state of the enzyme (347).

The position of the protomer-polymer equilibrium can be shifted toward the protomer by converting the enzyme to its carboxylated form, enzyme–biotin–CO_2^- (119,402) or by the binding of certain hydrophobic compounds, that is, long-chain fatty acyl CoA derivatives (474,477) or fatty acids (478-480). By introducing a carboxyl function at the 1'-N-position of the biotinyl prosthetic group, sufficient conformational strain is produced so that only true activators—that is, citrate and isocitrate, but not tricarballylate, P_i, or malonate—are able to constrain the enzyme and prevent its depolymerization (Section IV.B.2) (119,481). Malonyl CoA, which can carboxylate the enzyme, is a competitive inhibitor of the overall forward reaction (K_i = about 10^{-5} M) (402,431) with respect to citrate (402). This competitive kinetic relationship apparently results from the opposing effects of citrate and malonyl CoA on the protomer-polymer equilibrium. Like inhibition by malonyl CoA, that by long-chain acyl CoA derivatives ($K_i = 3-8 \times 10^{-7}$ M) (89,445,474) and fatty acids (480) can be explained by their effects on this equilibrium. Fatty acyl CoA derivatives, such as palmitoyl CoA, stearoyl CoA, and oleoyl CoA, are also competitive inhibitors with respect to citrate (89,474). Furthermore, these derivatives promote depolymerization of the carboxylase, which can be prevented by citrate (474). Reversal of the inhibition of acetyl CoA carboxylase by fatty acyl CoA derivatives with (+)-palmitoyl carnitine or the cetyl trimethylammonium ion is probably attributable to their displacement from the enzyme (482). There is abundant evidence indicating that a natural hydrophobic inhibitor of acetyl CoA carboxylase is present in crude enzyme extracts of liver and adipose tissue (136,407,482-485). The activating effect of (+)-palmitoyl carnitine or serum albumin on impure acetyl CoA carboxylase preparations (407,416,482-485) may also be due to the displacement of hydrophobic inhibitors, since homogeneous carboxylase preparations are unaffected by these substances (407,416).

Significant differences exist among acetyl CoA carboxylases from different sources with respect to the rate of the activation and polymerization processes   (89,131,388,401,407,408,474-476,482,486,487). The avian liver and bovine adipose tissue carboxylases are instantaneously activated (<10 sec) by tricarboxylic acid effectors (401,407), whereas the enzymes from rat liver, mammary gland, and adipose tissue necessitate preliminary incubation at 25-38° for 15-30 min in

$$\text{Protomer} \underset{\substack{+\text{Malonyl CoA} \\ \text{or} \\ +\text{ATP, Mg}^{2+}\text{ and} \\ \text{HCO}_3^-}}{\overset{+\text{Citrate}}{\rightleftharpoons}} \text{Polymer (filamentous)}$$

Characteristics	Protomer	Polymer (filamentous)
Catalytic activity	Essentially inactive	Active
$s_{20,w}$	13.1 S	47-59 S
Molecular weight	410,000	4-10 million

Fig. 14. Protomer-polymer transitions of avian liver acetyl CoA carboxylase.
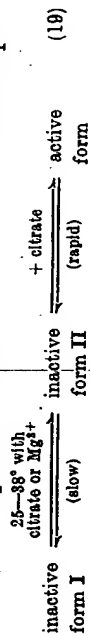
microscopy (401,406), and a large increase in intrinsic viscosity (347). Furthermore, the close correlation of the kinetics of depolymerization of the filamentous form with the rate of loss of enzymatic activity

order to achieve full activity when subsequently assayed with citrate (89,131,388,408,474–476,482,486,487). Rat-liver acetyl CoA carboxylase has been shown to be reversibly inactivated by low temperature (0°) and reactivated by preliminary incubation with tricarboxylic acid activator at 25° prior to assay (89,475). According to Greenspan and Lowenstein (476,482), activation can also be achieved by preliminary incubation at 37° with magnesium ion; however, citrate is still required as activator in the subsequent assay. These findings suggest that two conformational transitions may be necessary for activation with the carboxylase from rat liver: (a) a slow temperature-dependent process requiring magnesium ion (476,482,486,487) or citrate (89,131, 388,408,474,475,489), and (b) a rapid citrate-dependent process (mechanism 19). It is possible that the avian liver and bovine adipose

$$\text{inactive form I} \underset{\text{(slow)}}{\overset{\substack{25-38° \text{ with} \\ \text{citrate or Mg}^{2+}}}{\rightleftharpoons}} \text{inactive form II} \underset{\text{(rapid)}}{\overset{+ \text{ citrate}}{\rightleftharpoons}} \text{active form} \quad (19)$$

tissue carboxylases undergo similar transitions, but that both processes are rapid and, therefore, not readily detected.

Another activation phenomenon, also related to the state of polymerization of the carboxylase from rat liver, is the activation of the enzyme by trypsin treatment (490). It was first demonstrated by Anker and co-workers (488,490) and later in Numa's laboratory (93,491) that the rat-liver carboxylase could be activated by incubation with trypsin at 0° or 25° when assayed subsequently at 25-33°. This activation, which was 30–50% of that citrate with untreated enzyme, was accompanied by polymerization of the carboxylase at 25° (93), but not at 5° (488). Thus it would appear that limited peptide-bond cleavage of the carboxylase by trypsin gives rise to a form of the enzyme which exhibits catalytic and aggregational properties characteristic of those of the citrate-activated enzyme (93).

b. Mechanism of Citrate Activation. Initial investigations into the mechanism of activation by tricarboxylic acids revealed that the principal kinetic effect of the activator is on the V_{max} and that the K_m values for the substrates (ATP, HCO_3^-, and acetyl CoA) are not materially affected (136,407,474). The results of direct binding indicate that the affinity of the avian liver carboxylase for acetyl CoA is not affected by citrate (119). Thus it appears that activation increases the rate of reaction of bound substrate, rather than altering the affinity of the enzyme for substrates, and therefore would be classified as a positive allosteric effector of the V_{max} type (492).

Of importance in this regard is the finding that both partial reactions (reactions 17a and b), which are involved in the overall carboxylation, are activated by citrate and isocitrate. The fact that the exchanges which characterize the first partial reaction have no acetyl CoA requirement and that which characterizes the second partial reaction has no ATP, ADP, P_i, Mg^{2+}, or HCO_3^- requirement, exhibit citrate activation supports the view that a V_{max} rather than a K_m effect is of primary kinetic importance in the activation mechanism (139). Studies with model partial reactions, that is, the carboxylation of free biotin and carboxyl transfer from Enz–CO_2^- to acetyl pantetheine (an acetyl-CoA analogue) also show that both partial reactions are citrate-activated (140). These results suggest that a substituent at the active site of the enzyme, presumably the biotinyl prosthetic group, common to and involved in the rate-limiting steps of both half-reactions, is the focal point of the citrate-induced conformational change. A conformational change in the vicinity of the biotinyl group promoted by citrate is reflected by its inaccessibility to avidin (Section VII.A), the biotin-binding protein from egg white (98,139). Apparently the prosthetic group becomes shielded by neighboring groups as a result of conformational changes at the active site induced by citrate. The proposal that the biotin prosthetic group is the focal point of citrate-induced conformational changes is supported by the observation that the reactivity of the carboxylated prosthetic group toward decarboxylation (reaction 20) is enhanced by tricarboxylic acid activators, particularly in the presence of acetyl CoA (139,311,423,425). At pH

$$\text{Enz–biotin–}CO_2^- \rightarrow \text{Enz–biotin} + CO_2 \quad (20)$$
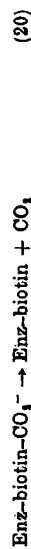
7.5 and 25° the first-order decarboxylation rate for carboxybiotinyl enzyme is increased 5-fold by citrate and ≥20-fold by citrate in the presence of less than saturating concentrations of acetyl CoA (Table V). The activator specificity patterns and other parameters for enzyme–CO_2^- decarboxylation are similar to those for the overall carboxylation (136,139). The avidin-insensitive decarboxylation of malonyl CoA catalyzed by the avian liver enzyme (Section IV.C.1) (347) is not activated by citrate, further suggesting that the biotinyl group participates in the activator-stimulated decarboxylation. The reactivity of the carboxybiotin prosthetic group, as measured by its susceptibility to decarboxylation, appears to reflect its reactivity with respect to carboxyl transfer to acetyl CoA. (Section IV.B.2).

Since the biotinyl prosthetic group is located at the end of a flexible
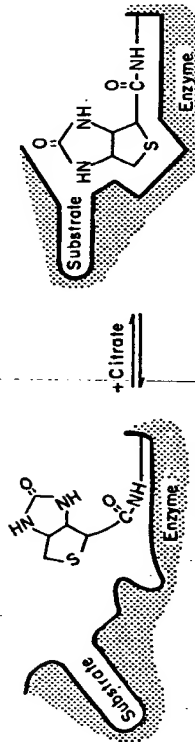
Fig. 15. Hypothetical scheme for the reorientation of the active site in the presence of tricarboxylic acid activator. From Lane et al. (311).

14-A side chain which is attached to the apoprotein, the precise orientation of the ureido ring system is required for its interaction with bound substrate. One possible mechanism for the citrate effect on transcarboxylation rate, supported by recent investigations (311), is that the carboxybiotin prosthetic group may be brought into closer proximity to the substrate-binding sites by activator-induced conformational changes (Fig. 15). Furthermore, citrate may enhance the affinity of a binding site in the vicinity of the substrate for the bicyclic ring of the prosthetic group and thereby facilitate reaction. This would account for the V_{max} effect of the tricarboxylic acid activator, as well as all other experimental findings associated with the activation phenomenon.

B. PROPIONYL CoA CARBOXYLASE

1. Background

Propionate arises from the oxidation of odd-numbered fatty acids (493–495), catabolism of branched-chain amino acids (496–499), and fermentation by the gastrointestinal flora of ruminants and other animals (500–504). This 3-carbon acid, which is actively metabolized by animal tissues (505,506), was shown as early as 1912 by Ringer to be a glycogenic compound (507). The glycogenic nature of propionate (506,508–510) was initially attributed to its conversion to pyruvate through acrylate and lactate (511–513). However, studies by Lorber et al. (514) and others (515,516) with propionate-2 or 3-^{14}C showed that complete randomization of label occurred during its conversion to glycogen, acetyl groups, or lactate. Work in Lardy's laboratory (86, 517,518) demonstrated that extracts of acetone-dried mitochondria catalyzed the ATP- and divalent cation-dependent carboxylation of propionate to form succinate. In addition, it was observed that this carboxylating activity was markedly depressed in the liver mitochondrial extracts from biotin-deficient rats (86), providing the first

indication that biotin was involved in propionate metabolism. Shortly thereafter, Flavin and Ochoa (104,106,519) and Katz and Chaikoff (105) found that the apparent initial product of propionate carboxylation was methylmalonate rather than succinate. Furthermore, it was demonstrated that propionate was first converted to propionyl CoA and was then carboxylated to form methylmalonyl CoA (106,236,519). It was discovered in Ochoa's laboratory (520–523) that the isomerization of methylmalonyl CoA to succinyl CoA could be resolved into two steps: (a) racemization of the methylmalonyl CoA formed by the carboxylase to its optical isomer (523,524), followed by (b) intramolecular thioester carbonyl migration to the C-3 position which is catalyzed by a vitamin B$_{12}$ coenzyme-dependent mutase (525–530, 531–537). The currently accepted pathway for the metabolism of propionate in animal tissues is shown in Figure 16. Abnormalities in the metabolism of methylmalonyl CoA in humans, which lead to elevated levels of excretion of methylmalonate in the urine, have been reported (538–546). These symptoms have been shown to result from defects at the methylmalonyl CoA mutase-catalyzed step (543–545).

Propionyl CoA carboxylase has been isolated in crystalline form from pig heart (547), in essentially homogeneous form from bovine liver mitochondria (548), and in partially purified form from rat-liver mitochondria (549) and Mycobacterium smegmatis (181). The direct participation of biotin in the carboxylase-catalyzed reaction was demonstrated initially by the sensitivity of the enzyme to avidin

Fig. 16. Pathway for the metabolism of propionate in animal tissues.

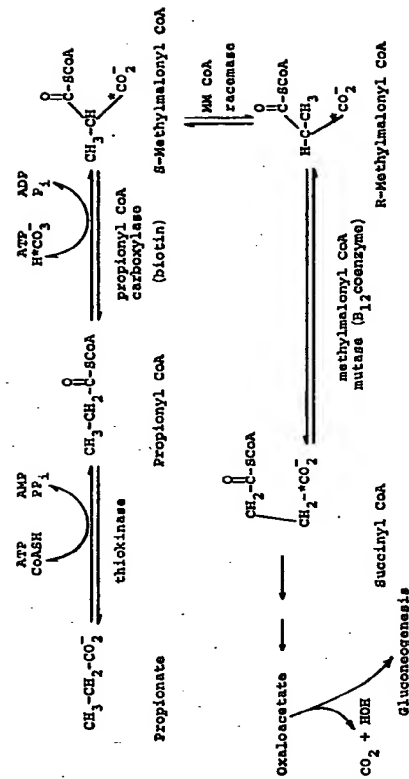(126,550) and later by the determination of the biotin content of the purified enzyme by microbiological assay (126,152).
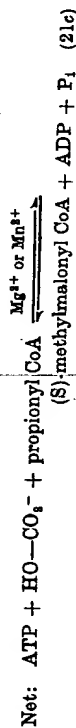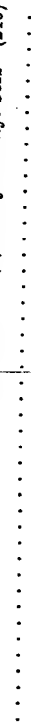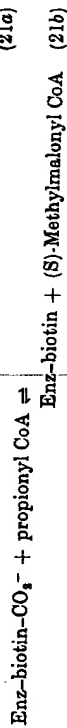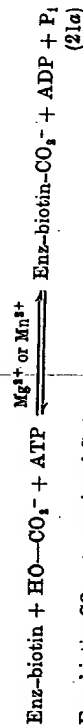
2. Reaction and Mechanism

The ATP- and Mg²⁺-dependent carboxylation of propionyl CoA to form (S)-methylmalonyl CoA, ADP, and P_i (reaction 21c) is a readily reversible reaction. At pH 8.1 and 28° the equilibrium constant

$$K' = \frac{(ADP)(P_i)((S)\text{-methylmalonyl CoA})}{(ATP)(HCO_3^-)(\text{propionyl CoA})}$$
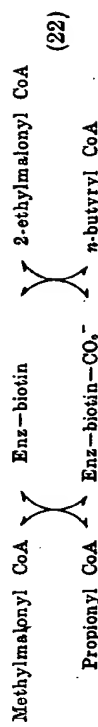
was found to be 5.7 which corresponds to a free energy change (ΔF_{301}) of -1.03 kcal/mole (182).

Propionyl CoA carboxylase catalyzes the minimal partial-reaction, sequence shown below. The first half-reaction was indicated by the

$$\text{Enz–biotin} + HO\text{—}CO_2^- + ATP \underset{}{\overset{Mg^{2+} \text{ or } Mn^{2+}}{\rightleftharpoons}} \text{Enz–biotin–}CO_2^- + ADP + P_i \quad (21a)$$

$$\text{Enz–biotin–}CO_2^- + \text{propionyl CoA} \rightleftharpoons \text{Enz–biotin} + (S)\text{-Methylmalonyl CoA} \quad (21b)$$

$$\text{Net: } ATP + HO\text{—}CO_2^- + \text{propionyl CoA} \underset{}{\overset{Mg^{2+} \text{ or } Mn^{2+}}{\rightleftharpoons}} (S)\text{-methylmalonyl CoA} + ADP + P_i \quad (21c)$$

ability of the enzyme to catalyze ATP–³²Pᵢ, and ATP–³²P-labeled ADP exchange dependent on all components of reaction 21a, but independent of those of the second half-reaction (reaction 21b) (126). Although the ATP–³²Pᵢ exchange rate was slower than anticipated (126,129) when the competition between ATP and ADP for a common binding site is taken into consideration, the exchange rate is compatible with that of the overall forward and reverse reactions (129,551). As pointed out earlier (Section I), substrate quantities of a carboxylated biotin enzyme (enzyme–CO_2^-) were first isolated by Kaziro and Ocha (83) following a short incubation of propionyl CoA carboxylase with ATP, bicarbonate, and divalent cation. Decarboxylation of isolated enzyme–CO_2^- in the presence of ADP, P_i, and Mg²⁺, as well as quantitative transfer of the "labile" carboxyl group to propionyl CoA, substantiates the formation of an active enzyme–CO_2^- intermediate (83). The site of carboxylation of the enzyme–CO_2^- intermediate was shown subsequently by Lane and Lynen (84) to be the $1'$-N position of the biotinyl prosthetic group (Section IV.B.1). Indirect evidence for the occurrence of the first partial reaction was obtained from the slow

bicarbonate- and magnesium-dependent cleavage of ATP to ADP and P_i catalyzed by propionyl CoA carboxylase. This ATPase activity is accounted for by the rapid formation of enzyme–CO_2^- via the first partial reaction (reaction 21a), followed by its slow spontaneous decarboxylation (Section IV.B.2) (83). The ATPase activity is enhanced by the addition of arsenate and ADP (129). As discussed earlier (Section II.B.2a), reversal of the first partial reaction (reaction 9b) results in the formation of an unstable "ADP-arsenate" intermediate which undergoes rapid nonenzymatic hydrolysis (reaction 9c).

The occurrence of the second partial reaction (reaction 21b) was first indicated by the demonstration that propionyl CoA carboxylase catalyzes an exchange between methylmalonyl CoA and ¹⁴C-propionyl CoA which is independent of the components of the first partial reaction (reaction 21a) (129,148,149,152). Further support for this half-reaction was provided by the observation that net carboxyl transfer from methylmalonyl CoA to n-butyryl CoA forming propionyl CoA and 2-ethylmalonyl CoA (reaction 22) is also catalyzed by the carboxylase in the absence of ATP, ADP, P_i, HCO_3^-, and divalent metal ion (147,150). In view of the fact that acetyl CoA can substitute for either

$$\text{Methylmalonyl CoA} \quad \text{Enz–biotin} \quad \text{2-ethylmalonyl CoA}$$
$$\diagdown \diagup \qquad \diagdown \diagup$$
$$\text{Propionyl CoA} \quad \text{Enz–biotin–}CO_2^- \quad \text{n-butyryl CoA} \quad (22)$$

propionyl CoA or n-butyryl CoA, it is evident that transcarboxylation can occur within the limits of the substrate specificity of the carboxylase (147,150,550,552). Both of these reactions implicated the intermediate participation of enzyme–CO_2^-. Direct formation of enzyme–¹⁴CO_2^- from ¹⁴C-carboxyl-labeled methylmalonyl CoA was demonstrated subsequently by Kaziro and Ochoa (83).

The dependence of methylmalonyl CoA–H¹⁴CO_3^- exchange on ADP, P_i, and divalent metal ion (107,147) indicates that reversal of the overall reaction involving both partial reactions is required for the incorporation of H¹⁴CO_3^- into methylmalonyl CoA. The fact that this reaction, as well as all of the reactions above, is blocked by avidin (83,107,126,129,149,150,152) is consistent with the role of the biotin prosthetic group as a carboxyl carrier.

Several lines of evidence supporting a concerted mechanism of carboxybiotinyl enzyme formation via the first partial reaction (reaction 21a) for acyl CoA carboxylases in general have been obtained with propionyl CoA carboxylase. These include (a) the dependence of

ATP-^{32}P$_i$ and ATP-^{32}P-labeled ADP exchange on all components of the first partial reaction (126,129), (b) the dependence of methyl-malonyl CoA–H^{14}CO$_3^-$ exchange on ADP, P$_i$, and divalent metal ion (107,147), and (c) the dependence of the arsenate-stimulated ATPase activity on ADP, HCO$_3^-$, and divalent metal ion (129). As pointed out in Section IV.A, alternative stepwise mechanisms involving the ordered release of ADP and P$_i$ are also compatible with the above findings. Nevertheless, any mechanism proposed must account for the incorporation of ^{18}O, from HC^{18}O$_3^-$, into P$_i$ arising from the γ-phosphoryl group of ATP (129). For further discussion of this subject, see Figure 6 and Section IV.A. The fact that the ratio of ^{18}O from HC^{18}O$_3^-$ incorporated into the carboxyl of methylmalonyl CoA to that incorporated into P$_i$ was approximately 2, indicates that HCO$_3^-$, rather than CO$_2$, is the carboxylating species (129). Bicarbonate was found subsequently by an independent method to be the "active species" of pyruvate carboxylase, another biotin-dependent enzyme (295,296).

The carboxylation of propionyl CoA, in the second partial reaction, occurs with retention of configuration at the C-2 position, giving rise to (S)-methylmalonyl CoA, indicating that proton abstraction and carboxyl addition occur from the same side of the C-2 carbon atom (339,343–345). Furthermore, tritium release from the C-2 position takes place at a rate equal to that of propionyl CoA carboxylation and occurs without a tritium isotope effect (339). As discussed previously (Section IV.C.1), these results are consistent with a mechanism in which carboxylation occurs in concert with proton abstraction at C-2.

In addition to catalyzing the carboxylation of propionyl CoA, the enzyme also catalyzes the much slower carboxylation of acetyl CoA, n-butyryl CoA, isobutyryl CoA, valeryl CoA, and crotonyl CoA to form their corresponding malonyl CoA derivatives (107,553–555). The fact that the maximal velocities of carboxylation of these CoA thioesters differ by two orders of magnitude, while their K_m values are nearly the same (152) suggests that the CoA moiety is of primary importance in substrate binding and that the acyl portion mainly influences the reaction velocity. In agreement with this proposal are the findings that valeryl CoA, CoA, and 3'-AMP are competitive inhibitors with respect to propionyl CoA and that removal of the 3'-phosphoryl and the adenine nucleotide groups from the thioester greatly increases the K_m value of the substrate (152,548). It appears, therefore, that the 3'-phosphate, adenine, and pantoyl moieties are involved in binding the acyl CoA substrate to the carboxylase. An

enzyme sulfhydryl group was implicated in the binding of acyl CoA substrate at the active site, since propionyl CoA dramatically protects the enzyme from inhibition by p-chloromercuribenzoate (p-CMB) (548). It was subsequently found by Edwards and Keech (556,557) that one sulfhydryl group, contributed by a cysteinyl residue of the enzyme, is specifically involved in binding propionyl CoA. Acyl CoA substrates do not appear to bind covalently, that is, by formation of acyl-(S)-enzyme as in the case of fatty acid synthetase (374,558–563), with the release of free CoA. Incubation of purified propionyl CoA carboxylase with propionyl-1-^{14}C-CoA or methylmalonyl-3-^{14}C-CoA in the presence of high concentrations of pantetheine did not lead to the formation of either propionyl-1-^{14}C-pantetheine or methylmalonyl-3-^{14}C-pantetheine (152). It should be pointed out that propionyl pantetheine serves as a substrate ($K_m = 24$ mM) for the carboxylase (152).

Propionyl CoA carboxylases from pig heart (564,565), bovine liver (134,566), and rat liver (549) are activated by certain monovalent cations. Edwards and Keech (565) demonstrated with the enzyme from pig heart that potassium ion reduced the K_m for bicarbonate from 8 to 3 mM. In contrast, Giorgio and Plaut, working with the bovine liver carboxylase (134), did not observe an altered K_m for HCO$_3^-$, Mg^{2+}, or propionyl CoA. Rather, the rates of the exchange reactions associated with the first (ATP-P, and ATP-ADP exchange) and second (methylmalonyl CoA·propionyl-2-^{14}C-CoA exchange) partial reactions were increased.

3. Molecular Characteristics

Preparations of propionyl CoA carboxylase of sufficient purity for physical-chemical characterization have been obtained from pig heart (547), and from bovine liver mitochondria (152,548). When comparison is possible, these enzymes have been found to have similar molecular properties, for example, for the pig heart and bovine liver carboxylases, respectively, $s^0_{20,w} = 19.7 \pm 0.09$ S and $s_{20,w} = 19.0$ S (3 mg protein/ml); biotin content, 1.83 μg/mg of protein and 1.91* μg/mg of protein. Since the pig heart enzyme has a molecular weight (M_{SD}) of 697,000 and contains 1 mole of biotin per 175,000 grams of protein, there appear to be 4 biotinyl prosthetic groups per molecule of enzyme (547). It is

* Calculated from the biotin content (1.34 μg/mg of protein) for a preparation of 70% purity.

TABLE VI
Physical Properties of Propionyl CoA Carboxylase from Pig Heart (547)

Molecular weight: $M_{SD} = 679,000$
Sedimentation
coefficient: $s_{20,w}^0 = 19.72 \pm 0.09$ S

$$s_{20,w} = s_{20,w}^0 - kc$$
$k = 1.85$ S g^{-1} dl
c = concentration, g/100 ml

Diffusion
coefficient: $D_{20,w}^0 = (2.58 \pm 0.06) \times 10^{-7}$ cm^2/sec
$$D_{20,w} = D_{20,w}^0 - kc$$
$k = 1.05 \times 10^{-7}$ cm^2 sec^{-1} g^{-1} dl
δ = mean protein concentration, g/100 ml

Intrinsic
viscosity: $[\eta] = 0.083$ dl/g
f/f_0: 1.41; axial ratio = 6.2–12.5 by different methods
Isoelectric
point: 6.1
Crystal form: hexagonal prisms with pyramidal bases.

interesting that several other biotin-dependent carboxylases—pyruvate carboxylase from liver (120) and yeast (567–569) and β-methylcrotonyl CoA carboxylase (570)—also contain 4 biotinyl prosthetic groups per molecule. This suggests that propionyl CoA carboxylase may have a tetrameric structure consisting of 4 protomeric units similar to that of pyruvate carboxylase (Section VI.E.). Propionyl CoA carboxylase has been dissociated with 7 M urea into subunits having a sedimentation coefficient of 2.5 S (547). Dissociation is accompanied by loss of enzymatic activity (547). The molecular properties of the pig heart propionyl CoA carboxylase are summarized in Table VI.

C. β-METHYLCROTONYL CoA CARBOXYLASE

1. Background

The role of β-methylcrotonyl CoA carboxylase in the catabolism of leucine, a ketogenic amino acid (571–576), is well established (577–580). Isotope studies carried out in several laboratories (581–584) demonstrated that the degradation of leucine leads to the formation of isovalerate (581), acetoacetate (583,584), and acetate (582). ^{14}C-Bicarbonate was incorporated into the carboxyl group of acetoacetate

(585–587), the C-2, C-3, and C-4 carbon atoms arising from the terminal isopropyl group of leucine (584) as illustrated below and in Figure 17.

$$CH_3-\overset{CH_3}{\underset{CH_3}{C}}\!\!\begin{bmatrix}CH-CH_2-CH-CO_2H \\ NH_2\end{bmatrix}$$
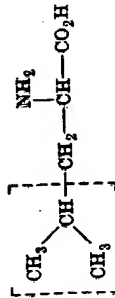
The C-2 and C-3 positions of leucine (582) and the C-1 and C-2 positions of isovalerate (584) gave rise to a two-carbon intermediate which can condense to form acetoacetate. Coon et al. (408) proposed that isovalerate was converted to its CoA thioester and subsequently dehydrogenated to form β-methylcrotonyl CoA. Furthermore, it was demonstrated in both bacterial (588,589) and animal systems (590) that β-methylcrotonyl CoA was a precursor of β-hydroxy-β-methyl-glutaryl CoA. Lynen et al. (80) showed conclusively with a purified bacterial carboxylase that β-methylcrotonyl CoA underwent an ATP-dependent carboxylation to form β-methylglutaconyl CoA. This observation was confirmed by Coon and co-workers (111,591,592) using
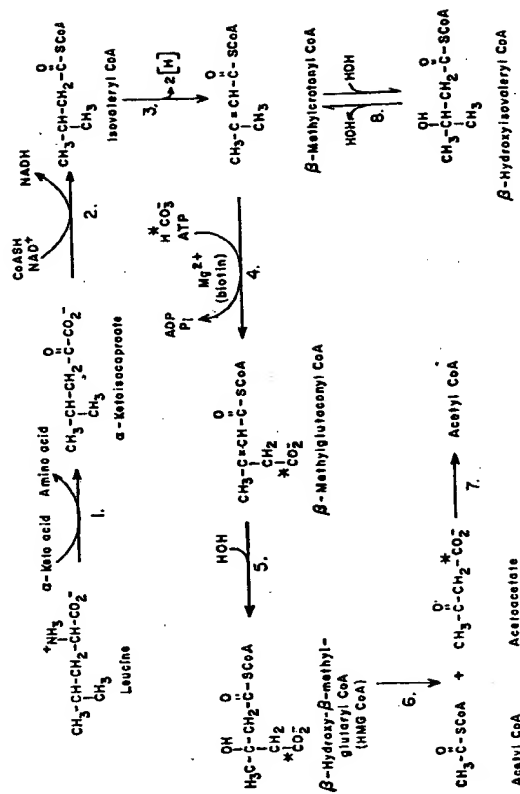
Fig. 17. Pathway for the catabolism of leucine.

a purified carboxylase, indicating that the substrate of the carboxylase was not β-hydroxyisovaleryl CoA as originally proposed (580,593–595). The formation of β-hydroxy-β-methylglutaryl CoA from β-methylglutaconyl CoA was shown to be catalyzed by a specific hydrase present in bacterial and animal extracts (589). The final step (reaction 6 in Fig. 17), that is, the cleavage of β-hydroxy-β-methylglutaryl CoA yielding acetyl CoA and acetoacetate, is accounted for by a specific cleavage enzyme characterized earlier by Bachhawat et al. (596). An inherited metabolic disorder, "maple syrup urine disease," characterized by a deficiency in oxidative decarboxylation of the α-keto acids derived from leucine and other branched-chain amino acids (reaction 2 in Fig. 17) has been demonstrated in humans (597).

Evidence for the participation of a biotin-dependent carboxylase in this pathway in animal tissues was indicated by (a) the inability to convert isovalerate to acetoacetate (598), (b) a reduced level of $H^{14}CO_3^-$ incorporation into acetoacetate (599), and (c) the depression of β-hydroxyvaleryl CoA-dependent bicarbonate fixation as a result of biotin deficiency (600).

2. Reaction and Mechanism

Much of our fundamental knowledge regarding the mechanism of action of the biotin-dependent enzymes emanated from investigations in Lynen's laboratory on the β-methylcrotonyl CoA carboxylase (80–82,138,577). The important discovery that this enzyme catalyzed the rapid ATP- and divalent cation–dependent carboxylation of free (+)-biotin (80) led to the identification of the carboxylation site as the 1'-N position of biotin (81,82) as discussed in detail in Section IV.B.1. It was reasoned that since the carboxylation 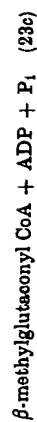reaction exhibited a high degree of structural and stereochemical specificity for free (+)-biotin, carboxylation probably occurred at the same enzyme site normally occupied by the bicyclic ring of the biotinyl prosthetic group (81). This implied that the bicyclic ring could be displaced from its site while remaining covalently attached to the enzyme through its side chain, permitting access to this site of the free derivative. The recent isolation and characterization of an essential component (biotin carboxylase) from E. coli acetyl CoA carboxylase, which is free of biotin and catalyzes the carboxylation of free biotin, validates this proposal (Section IV.B.1) (95,96,100). Later investigations by Knappe and co-workers (86) showed that, like free biotin, the biotin prosthetic

group of β-methylcrotonyl CoA is carboxylated at the 1'-N position (Section IV.B.1).

Like the other biotin-dependent carboxylases, β-methylcrotonyl CoA carboxylase catalyzes the avidin-sensitive (80,81) 2-step reaction sequence shown below. Consistent with the occurrence of the first

$$\text{Enz–biotin} + \text{HO—CO}_2^- + \text{ATP} \xrightarrow{\text{Mg}^{2+} \text{ or } \text{Mn}^{2+}} \text{Enz–biotin–CO}_2^- + \text{ADP} + P_i \qquad (23a)$$

$$\text{Enz–biotin–CO}_2^- + \beta\text{-methylcrotonyl CoA} \rightleftharpoons$$
$$\text{Enz–biotin} + \beta\text{-methylglutaconyl CoA} \qquad (23b)$$

$$\text{Net:} \quad \text{ATP} + \text{HO—CO}_2^- + \beta\text{-methylcrotonyl CoA} \xrightarrow{\text{Mg}^{2+} \text{ or } \text{Mn}^{2+}}$$
$$\beta\text{-methylglutaconyl CoA} + \text{ADP} + P_i \qquad (23c)$$
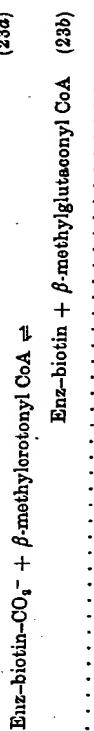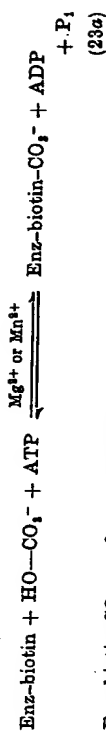
half-reaction (reaction 23a) was the finding of Lynen et al. (81) that the carboxylase from Achromobacter catalyzes exchanges between ATP and $^{32}P_i$ and between ATP and ^{14}C-ADP in the absence of acyl CoA substrates. The dependence of both of these exchange reactions on ATP, ADP, HCO_3^-, P_i, and divalent metal ion (81) favors a concerted mechanism for the carboxylation of the biotin prosthetic group, although unequivocal evidence on this point has not been obtained with any biotin-dependent carboxylase (Section IV.A). The ability of β-methylcrotonyl CoA carboxylases from Mycobacterium spp. (81) and Achromobacter (138) to catalyze the carboxylation of free biotin in the presence of ATP and Mg^{2+} provided an enzymatic model for the first step of overall reaction and precedent for the carboxylation of the prosthetic group. Himes et al. (138) and Knappe et al. (86) isolated the carboxylated form of β-methylcrotonyl carboxylase by gel filtration following a prior incubation with ATP, $H^{14}CO_3^-$, and divalent cation. It was demonstrated that these enzyme-$^{14}CO_2$ preparations (a) contained equivalent amounts of ^{14}C-labeled carboxyl and biotin (138), (b) were rapidly decarboxylated at acid pH (138), and (c) transferred their labile carboxyl group stoichiometrically to β-methylcrotonyl CoA (86,138). Following methylation and proteolytic digestion the ^{14}C-labeled methoxycarbonyl group was located at the 1'-N position of the biotinyl moiety (Section IV.B.1) (86). These results supported the proposal that enzyme-biotin-CO_2^- formed via the first half-reaction serves as a carboxyl donor in the second partial reaction. The earlier

observation that the carboxylase preparation from *Mycobacterium spp.* which was contaminated by β-methylglutaconase catalyzed a β-hydroxy-β-methylglutaryl CoA–^{14}C-β-methylcrotonyl CoA exchange strongly indicated that the second partial reaction occurred (reaction 23b) (81). All of the preceding exchange reactions are inhibited by avidin (81).

In contrast to acetyl CoA and propionyl CoA carboxylases which carry out carboxylation at the α-position of the thioester, β-methylcrotonyl CoA carboxylase directs carboxylation to the γ-position (81). From the chemical point of view, carboxylation at the γ-position would be predicted since conjugation through the double bond displaces the site of electrophilic activation to the γ-position. It is interesting in this regard that crotonyl CoA and acetoacetyl CoA, presumably in its enolate form, are also carboxylated at 20% and 14%, respectively, the rate of the normal substrate (138). The carboxylation product of crotonyl CoA was identified as glutaconyl CoA (138). The carboxylase from *Achromobacter*, which was used for these studies, carboxylates free (+)-biotin at a rate 24 to 30% that of β-methylcrotonyl CoA (138), its presumed natural substrate.

3. Molecular Characteristics

β-Methylcrotonyl CoA carboxylase from *Achromobacter*, which is homogeneous by analysis in the analytical ultracentrifuge and by Tiselius electrophoresis, has been obtained in crystalline form (570). The enzyme has a molecular weight (M_{SD}) of 760,000 ($s_{20,w}^0 = 20.7$ S; $D_{20,w}^0 = 2.84 \times 10^{-7}$ cm^2/sec) and contains 1 mole of biotin per 192,000 g of protein, which is compatible with a tetrameric structure (570). It is interesting that several other biotin-dependent carboxylases, including liver and yeast pyruvate carboxylase (120,567–569) and pig heart propionyl CoA carboxylase (547), apparently have tetrameric structures. As predicted from the high glutamate and aspartate content, the enzyme has an unusually low isoelectric point in vicinity of 3.5.

D. GERANYL CoA CARBOXYLASE

The bacterial degradation of isoprenoid compounds, for example, farnesol, geraniol, and citronellol, is initiated by the oxidation of the alcohol function to a carboxylate, followed by conversion to their corresponding acyl CoA derivatives (601). Further degradation, as
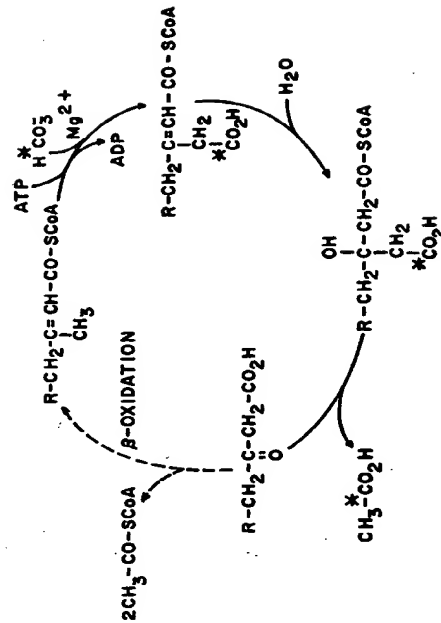
Fig. 18. Proposed pathway for the catabolism of geranyl CoA in *Pseudomonas citronellis*. From Seubert and Remberger (601).

visualized by Seubert and Remberger (601) in Figure 18, involves a biotin-dependent carboxylation reaction analogous to that catalyzed by β-methylcrotonyl CoA carboxylase. In the subsequent step the carboxymethyl group, containing the bicarbonate fixed in the preceding step, is eliminated, giving rise to a β-keto thioester. Following a β-keto thiolase-type reaction which generates acetyl CoA, the isoprenoyl CoA product can reenter the cycle.

A geranyl CoA carboxylase has been purified 200-fold from extracts of *Pseudomonas citronellolis* by Seubert et al. (112). The carboxylase catalyzes the ATP- and divalent metal ion–dependent carboxylation of geranyl CoA, farnesyl CoA, and β-methylcrotonyl CoA as shown below (reaction 24). Like the reaction catalyzed by β-methylcrotonyl CoA

$$\text{ATP} + \text{HO—CO}_2 + \text{R—CH}_2\text{—C}=\text{CH—C}=\text{CH—C—SCoA} \underset{\text{Mg}^{2+}}{\rightleftharpoons}$$

(with CH$_3$ and O)

$$\text{R—CH}_2\text{—C}=\text{CH—C—C—SCoA} + \text{ADP} + P_i \quad (24)$$

(with CH$_2$—CO$_2^-$ and O)
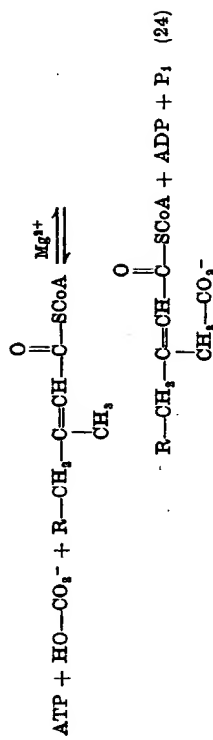
carboxylase (Section VI.C), the site of carboxylation is at the γ-methyl position. Among the thioester substrates tested, the enzyme

carboxylates geranyl CoA 3 and 18 times faster than farnesyl CoA and β-methylcrotonyl CoA, respectively, and is specific for the cis configuration (112). Since the carboxylase is completely blocked by avidin, but not by biotin-treated avidin, it is evident that the enzyme is biotin-dependent.

E. PYRUVATE CARBOXYLASE

1. Background

A carbon dioxide fixing reaction involving pyruvate was initially suggested by Wood and Werkman to explain the utilization of CO_2 by propionic acid bacteria grown on glycerol (602,609). On the basis of their observations that the stoichiometry of CO_2 fixed and succinate formed was 1 to 1 (607) and that pyruvate was found in the fermentation medium (606), it was proposed that oxaloacetate formed by the β-carboxylation of pyruvate was subsequently converted to succinate. Support for β-carboxylation was obtained by Krampitz, Wood, and Werkman (610) with an enzyme preparation from *Micrococcus lysodeikticus* which catalyzed an oxaloacetate (β-carboxyl)–$H^{13}CO_3^-$ exchange. Subsequently, a large body of information accumulated which indicated that carboxylation of pyruvate to form oxaloacetate ("Wood–Werkman reaction") occurs in both bacterial (611–614) and animal (615–626) systems. The fermentation of glucose-3,4-^{14}C by *Aspergillus niger*, which gives rise to citrate, involves (*a*) symmetrical cleavage of the hexose yielding two C_3 units, (*b*) decarboxylation of one C_3 unit → a C_2 unit, (*c*) carboxylation of the other C_3 unit → a C_4 unit, and finally (*d*) condensation of the C_2 and C_4 units to form citrate (627). Experiments with $^{14}CO_2$ indicated that the C-6 carboxyl group of citrate (627,628) is derived from the β-carboxyl of oxaloacetate. Further investigations by Woronick and Johnson (629) with the *Aspergillus* system indicated that cell-free extracts catalyzed the ATP-dependent carboxylation of pyruvate to form oxaloacetate.

The observation that liver mitochondria can form phosphoenol-pyruvate (PEP) from pyruvate (630–632) or oxaloacetate, the latter via PEP carboxykinase (632–640), prompted Utter and Keech to search for a reaction which could generate oxaloacetate from pyruvate. In 1960 they reported that ammonium sulfate-fractionated extracts of mitochondria from bovine and avian liver catalyzed the ATP, Mg²⁺, and acetyl CoA-dependent carboxylation of pyruvate to form oxalo-acetate (108,641). It became evident that the pyruvate carboxylase

reaction was biotin-dependent, since it was inhibited by avidin, and this inhibitory effect could be prevented by prior treatment of the avidin with biotin (108,641,642). This finding provided a basis for the earlier observation that aspartate could partially substitute for biotin in the nutrition of many microorganisms (32,38). The role of acetyl CoA was found to be that of an activator, since it was only required in catalytic quantities and labeled acetyl CoA did not appear in the oxaloacetate formed (108); a transcarboxylation mechanism involving the intermediate formation of malonyl CoA was ruled out, since tritium release from ^3H-methyl-labeled acetyl CoA did not occur and malonyl CoA was a potent inhibitor ($K_i = 7.3 \times 10^{-5}$ M) of the reaction (643). Of the pyruvate carboxylases isolated from animal and microbial sources, those from avian (108,644), rat (645), and bovine (108) liver, sheep kidney (646), baker's yeast (153,647–649), Rhizopus (650), Chromatium (651), *Bacillus stearothermophilus* (261), *Streptococcus faecalis* (652), *Arthrobacter globiformis* (653), and *Bacillus coagulans* (654) are activated by acetyl CoA, whereas those from *Aspergillus niger* (655) and *Pseudomonas citronellolis* (127) are not.

It was discovered by Scrutton et al. (157), using magnetic resonance methods to study the divalent cation requirement of the reaction, that purified pyruvate carboxylase enhanced the proton relaxation rate of water. Subsequently these investigators showed that this effect was caused by tightly bound paramagnetic manganese in the enzyme (157, 656). The presence of this metal was established by chemical, neutron activation, emission spectral, and atomic absorption spectral analysis (157). The manganese content (4 gatoms per 600,000 g of protein) of pyruvate carboxylase was found to be equimolar with respect to biotin (157,657,658). Furthermore, ⁵⁴Mn administered to growing chicks was incorporated into liver mitochondrial pyruvate carboxylase and, in the isolated enzyme, did not exchange with unlabeled manganese (156). The enzyme-bound manganese can only be removed by drastic conditions which lead to irreversible inactivation of the enzyme (157). Recent studies by Grimiger and Scrutton (659) showed that pyruvate carboxylase, isolated from chicks grown on a manganese-deficient, magnesium-sufficient diet, contained magnesium in place of manganese and was fully active. Pyruvate carboxylases from other sources, that is, bovine and turkey liver, also contain tightly bound manganese (656), whereas that from baker's yeast is a zinc metallo-enzyme (336,568). Since all other pyruvate carboxylases investigated

are inhibited by oxalate (649,656,660), a potent metal-binding inhibitor of the avian liver carboxylase, it appears that they too are metalloenzymes.

2. Reaction and Mechanism

The pyruvate carboxylase-catalyzed reaction can be partitioned into at least two distinct steps:

$$\text{Enz–biotin} + \text{HO—CO}_2^- + \text{ATP} \xrightarrow[\text{acetyl CoA}]{\text{Mg}^{2+}} \text{Enz–biotin–CO}_2^- + \text{ADP} + \text{P}_i \quad (25a)$$

$$\text{Enz–biotin–CO}_2^- + \text{pyruvate} \underset{\text{(tightly bound Mn)}}{\rightleftharpoons} \text{Enz–biotin} + \text{oxaloacetate} \quad (25b)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\text{Net: } \text{HO—CO}_2^- + \text{ATP} + \text{pyruvate} \xrightarrow[\text{(tightly bound Mn)}]{\overset{\text{Mg}^{2+}}{\text{acetyl CoA}}} \text{oxaloacetate} + \text{ADP} + \text{P}_i \quad (25c)$$
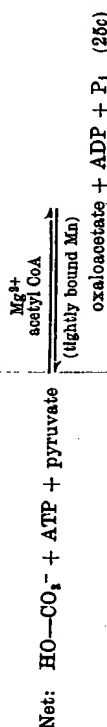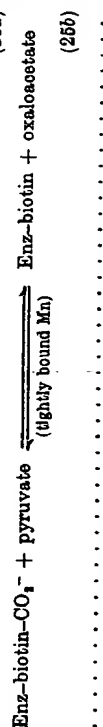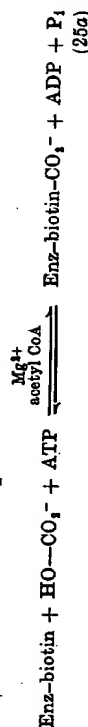
Precedent for these partial reactions was established by investigations with the acyl CoA carboxylases, that is, β-methylcrotonyl CoA carboxylase (Section VI.C) and propionyl CoA carboxylase (Section VI.B). As in the case of the propionyl CoA carboxylase-catalyzed reaction (129), the active species of "CO₂" involved in the pyruvate carboxylation mechanism was found to be HCO_3^- (Section IV.A) (296). Two unique features of the reaction catalyzed by pyruvate carboxylase are the activation of the first partial reaction (reaction 25a) by acetyl CoA (132) and the participation of tightly bound metal ion in the second partial reaction (reaction 25b) (155).

The isolation of a carboxylated enzyme intermediate, the formation of which is dependent on ATP, Mg^{2+}, and acetyl CoA and whose formation is inhibited by avidin, verifies the occurrence of the first partial reaction (132). Stoichiometric carboxyl transfer from enzyme–CO_2^- to pyruvate occurs in the absence of the components of the first partial reaction. The carboxylase catalyzes an ATP-$^{32}P_i$ exchange requiring HCO_3^-, Mg^{2+}, and ADP which is also absolutely dependent on acetyl CoA in the case of the avian liver enzyme (132), is stimulated by acetyl CoA with the yeast enzyme (661), and is independent of acetyl CoA with the P. citronellolis enzyme (127). This exchange reaction is avidin-sensitive and occurs at a rate about 0.5% that of the overall carboxylation reaction (127,132,137,661). While it has been

proposed that the first partial reaction may proceed via a concerted mechanism in the case of the acyl CoA carboxylases (Section IV.A), the ability of pyruvate carboxylase to catalyze an avidin-insensitive ATP-^{14}C-ADP exchange independent of P_i and requiring only Mg^{2+} raises some doubt as to the concertedness of this step (137). The similarity of the inhibition pattern by nucleoside triphosphates of this exchange and the overall reaction indicates that both are catalyzed by pyruvate carboxylase (137). These observations can be interpreted in terms of a phosphoryl-enzyme intermediate as discussed earlier (Section IV.A). The suggestion by Lynen (136) that biotin per se may be the site of phosphorylation is not inconsistent with the observation that ATP-^{14}C-ADP exchange is inhibited by avidin. It is not clear whether the proposed intermediate is on the main reaction pathway or is the result of a side reaction. Like the ATP-$^{32}P_i$ exchange, the rate of the ATP-^{14}C-ADP exchange is slow relative to the overall forward reaction (0.2–0.4% of the rate of pyruvate carboxylation) (137). The significance of these slow rates of exchange remains to be explained.

The participation of the second partial reaction (reaction 25b) in the overall process is supported by the occurrence of a carboxylase-catalyzed oxaloacetate-^{14}C-pyruvate exchange which is independent of the components of the first partial reaction (127,132,153). The second step is characterized by the lack of an acetyl CoA requirement (153) and the participation of tightly bound metal (155). With the avian liver carboxylase neither oxaloacetate-^{14}C-pyruvate exchange nor stoichiometric carboxyl transfer from enzyme–CO_2^- to pyruvate requires acetyl CoA (132). Like the overall carboxylation reaction, oxaloacetate-^{14}C-pyruvate exchange is inhibited by oxalate, a metal complexing agent, whereas the exchanges characteristic of the first partial reaction are not (155).

The role of the tightly bound metal ion in the second partial reaction has been most rigorously investigated with the manganese-containing avian liver carboxylase (155,157). The valence state of the bound manganese appears to be $+2$, according to magnetic susceptibility measurements (157). Studies on the role of the metal have been facilitated by the use of approaches relying on the paramagnetic properties of the bound manganese. It has been demonstrated by Mildvan et al. (155), using pulsed nuclear magnetic resonance techniques, that the bound manganese enhances ($\epsilon_b = 4.2$) the longitudinal

proton relaxation rate (PRR) of water about fourfold. Substrates of the second partial reaction, that is, pyruvate and oxaloacetate, cause a decrease in the enhanced PRR owing to bound manganese, whereas the substrates of the first reaction have no effect. The decreased enhancements (ϵ_0 = 1.7 for pyruvate, ϵ_0 = 1.9 for oxaloacetate) indicate the existence of complexes of pyruvate carboxylase with either pyruvate or oxaloacetate (662). A stable complex between the pyruvate carboxylase and oxalate, which blocks the second partial reaction, is indicated by the large decrease in enhancement (to ϵ_0 = 0.3) caused by oxalate (115). Evidence favoring a catalytically productive role for the complexes with pyruvate and oxaloacetate includes the observation that kinetic constants for these ligands obtained from initial velocity, inhibition, and avidin inactivation studies are in reasonably good agreement with those determined by the PRR method (662,663). Conclusive evidence in the case of pyruvate (337), and suggestive evidence, in the case of oxaloacetate (664), for an enzyme–manganese–keto acid bridge complex has been obtained by examining the effect of enzyme-bound manganese on the relaxation rates of the methyl and methylene protons of the respective substrates. It has been visualized (662) as shown in Figure 10 that pyruvate and oxaloacetate are mono- and bi-dentate ligands, respectively. In the case of the carboxylase–pyruvate complex the distance between the methyl protons and enzyme-bound manganese was calculated from the longitudinal proton relaxation rate to be 3.5 ± 1 Å (337) and for the methylene protons in the carboxylase–oxaloacetate complex, 4.6 ± 3 Å (664,665). These distances are compatible with the direct coordination of pyruvate and oxaloacetate with enzyme-bound manganese (664,666). The maximal exchange rates for pyruvate (337) and oxaloacetate (664) from the coordination sphere of the enzyme-bound manganese are at least 2 orders of magnitude greater than the turnover number of the catalytic site. Therefore the rate of formation and dissociation of the enzyme–manganese–substrate bridge complexes is sufficiently rapid to account for their participation in the catalytic mechanism.

It has been suggested (155) that the tightly bound metal ion of pyruvate carboxylase could function in at least two ways: (a) to bind and orient the keto acid substrate, and (b) to facilitate the removal of a methyl proton of pyruvate or the β-carboxyl group of oxaloacetate as visualized by the concerted mechanisms in Figure 10. As pointed out earlier (Section IV.C.2), recent studies (359) showing retention of

configuration at C-3 of pyruvate in the pyruvate carboxylase-catalyzed reaction are consistent with this mechanism.

3. Molecular Characteristics

Avian liver pyruvate carboxylase which has a molecular weight of 660,000 ($s_{20,w}^0$ = 14.8 S) appears to be composed of 4 protomers each having a molecular weight of 165,000 ($s_{20,w}^0$ = 6.75 S) (120,657,667). This is supported by its tetrameric appearance in the electron microscope (120,657). As shown in Figure 19, the negatively stained enzyme consists of 4 subunits arranged at the corners of squares (cyclic symmetry, Klotz et al. (668)) with the distance between the midpoints of adjacent subunits being 70 to 75 Å. This tetrameric structure is compatible with the finding that the enzyme contains approximately 4 biotin prosthetic groups (667) and 4 atoms of manganese (157) per carboxylase molecule. The protomeric units can be further dissociated into subunits of 45,000 daltons ($s_{20,w}^0$ = 2.7 S) with 1% sodium dodecyl sulfate (667). Pyruvate carboxylases from baker's yeast (567,569), turkey liver (567,656), and bovine liver mitochondria (567) also appear to have tetrameric structures in the electron microscope. While the appearance of the latter two carboxylases is similar (cyclic symmetry) to that of the chicken-liver enzyme, as illustrated in Figure 19, the
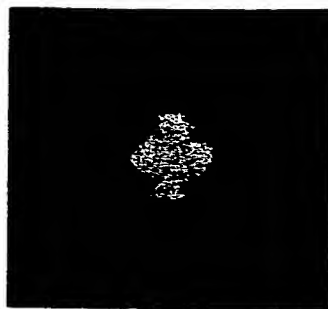

Fig. 19. Electron micrographs of pyruvate carboxylases from avian liver and yeast. Micrographs are superimpositions of 6 and 8 images, respectively, at magnifications of 1.25 × 10⁶ and 1.0 × 10⁶, respectively. The original images were obtained by negative staining with silicotungstate. From Young et al. (569).

yeast enzyme appears as 4 protomers arranged at the corners of a rhombus (657) (dihedral symmetry, Klotz et al. (668)). The pyruvate carboxylase from baker's yeast has a molecular weight of 600,000 ($s_{20,w}$ = 15.8 S) and contains approximately 4 atoms of zinc (338) and 4 biotinyl prosthetic groups (569) per molecule. Dissociation of the yeast enzyme to dimers ($s_{20,w}$ = 9–10 S) or monomers ($s_{20,w}$ = 6–7 S) can be accomplished with the controlled use of maleic anhydride (567, 569,656).

The chicken-liver pyruvate carboxylase has the interesting property of cold lability. Exposure to low temperature (2°) is accompanied by reversible loss of pyruvate carboxylase activity and reversible dissociation of the tetramer ($s_{20,w}$ = 14.8 S) to the protomeric form ($s_{20,w}$ = 6.75 S) (669). The dissociation phenomenon was verified by electron microscopy (120). In addition to the overall carboxylation reaction, the oxaloacetate-14C-pyruvate (132) and ATP-32Pi (132) exchanges are cold-labile, whereas the ATP-14C-ADP exchange is not (137). The exchange reactions are also restored by warming to 23° (132). Furthermore, cold inactivation leads neither to loss of tightly bound manganese, nor to a decreased enhancement of the proton relaxation rate owing to the bound metal ion (157). Acetyl CoA, an allosteric activator of the carboxylase, protects against the inactivation and dissociation resulting from exposure to low temperature (669), the acyl CoA specificity for this effect paralleling that for activation of the carboxylase (669). Protection from cold inactivation can also be achieved with glycerol, methanol, ATP, pyruvate, Mg^{2+}, HCO_3^-, and by increased protein concentration (656,667,669,670). Since cold inactivation accomplished in the presence of ATP leads to an inactive tetramer, Irias et al. (669) propose that protomer formation occurs via an inactive tetrameric intermediate. The rate-limiting step on subsequent reactivation is apparently the conversion of the inactive to the active tetramer. Prolonged exposure to low temperature apparently results in an irretrievable loss of enzymatic activity. Cold lability has also been demonstrated with the carboxylases from turkey liver (656), sheep kidney (646), and *A. globiformis* (653).

4. Effector Control

Pyruvate carboxylase from animal tissues exhibits an absolute requirement for catalytic concentrations of short-chain acyl CoA derivatives, acetyl CoA being the most effective (K_A = 2–3 µM for the avian liver enzyme) (656). This activation is instantaneous (<1 sec)

(671), reversible (643,671), and restricted to the first partial reaction (reaction 25a) (132). Furthermore, activation is not associated with a change in sedimentation velocity (671) as it is in the case of the acetyl CoA carboxylases from animal tissues (Section IV.A.4) (406). However, conformational changes in the vicinity of the prosthetic group have been demonstrated by the increased susceptibility of the biotinyl moiety to interaction with avidin accompanying activation by acetyl CoA (671). Activation by acetyl CoA and certain other short-chain acyl CoA derivatives (propionyl CoA, crotonyl CoA, and isobutyryl CoA) involves cooperative interactions and is characterized by Hill coefficients of about 2–3 (pH 7.8 avian liver enzyme) (656). Although the V_{max}s are similar, the activator constants for these derivatives differ by an order of magnitude, with that for acetyl CoA being lowest (656). The degree of cooperativity is pH-dependent, decreasing as pH increases (656). Competitive inhibitors with respect to acetyl CoA, for example, malonyl CoA (K_i = 0.47 mM), succinyl CoA (K_i = 0.43 mM), acetyl-3'-dephospho CoA (K_i = 1.2 mM), and oxidized CoA (K_i = 0.27 mM), do not alter the Hill coefficient for the activator (656).

A potential regulatory system involving acetoacetyl CoA and β-hydroxybutyryl CoA, which may complement activation by acetyl CoA, was reported recently (672). Acetoacetyl CoA significantly increases the apparent K_A and lowers the Hill coefficient (2.9 to 1) for acetyl CoA, while its reduced product, β-hydroxybutyryl CoA, is an activator. Fung and Utter (672) have suggested that the (NADH)/(NAD+) ratio which controls the (β-hydroxybutyryl CoA)/(acetoacetyl CoA) ratio may regulate the carboxylase through an effect on the K_A for acetyl CoA.

In contrast to the animal enzymes, pyruvate carboxylase from baker's yeast has a markedly different activation pattern with respect to CoA derivatives. Long-chain fatty acyl CoAs such as palmitoyl CoA (k_A = 0.2µM) are more effective activators than their short-chain counterparts (K_A = 80 µM for acetyl CoA) (567,656). Furthermore, the extent of activation with both of these effectors is approximately threefold in contrast to the absolute requirement for short-chain acyl CoA (656,673) derivatives with the animal carboxylases. Cooper and Benedict (674) have reported that the K_m for HCO_3^- is lowered by acetyl CoA. The yeast carboxylase is also activated by CoA (675) although there is disagreement regarding whether the activator constant is comparable to acetyl CoA (649) or several orders of magnitude greater than that for acetyl CoA. Activation by long-chain acyl CoA derivatives

would be compatible with an anaplerotic role (675,677) for the carboxylase at the time of fatty acid mobilization and oxidation.

The yeast carboxylase is also inhibited by L-aspartate ($K_i = 1.9$ mM) (628). Inhibition is noncompetitive with respect to pyruvate ATP, Mg²⁺, and HCO₃⁻ (678); however, it is antagonistic with respect to acetyl CoA, the cooperative response to acetyl CoA increasing in the presence of aspartate (679). The Hill coefficient rises from 1.1 in the absence of aspartate to 2.2 in the presence of 10 mM L-aspartate (679). Similar findings using palmitoyl CoA in place of acetyl CoA have been obtained in Utter's laboratory (656). L-Aspartate apparently serves as a negative feedback inhibitor on the carboxylase, in this way regulating the rate of its own synthesis. The K_A for L-aspartate is compatible with the finding that levels of L-aspartate in yeast cells fluctuate between 5 and 20 mM (680).

5. Gluconeogenic Role in Animal Tissues

Pyruvate carboxylase has been implicated in the formation of phosphoenolpyruvate (PEP) from pyruvate via a "dicarboxylic acid shuttle" (reactions 1 and 2, Fig. 20) which circumvents the thermodynamically unfavorable reversal of the pyruvate kinase-catalyzed reaction (reaction 3, Fig. 20) (640). This shuttle, which utilizes pyruvate carboxylase and PEP carboxykinase, is thought to be involved in gluconeogenesis from lactate and other precursors of pyruvate, for example, alanine, serine, and cysteine (681). In support of this proposed role for the "dicarboxylic acid shuttle," pyruvate carboxylase and PEP carboxykinase are localized primarily in gluconeogenic tissues such as liver and kidney (see Table II in reference 681) (682–685). In addition, the catalytic capacities of both enzymes are compatible with the rates of gluconeogenesis in these tissues (681). It is interesting that the hepatic intracellular localization of these two enzymes differs, pyruvate carboxylase being exclusively mitochondrial (642,681,686) whereas PEP carboxykinase may be found in the mitochondrial and/or extramitochondrial compartment depending on the species (640,681,687–691).

There is compelling evidence to indicate that pyruvate carboxylase, which functions at a crucial branch point in metabolism, has an important regulatory role in gluconeogenesis from pyruvate (108,641, 642,644–646,672,692–722). The activity of this enzyme fluctuates coordinately with the rate of gluconeogenesis induced by changes in
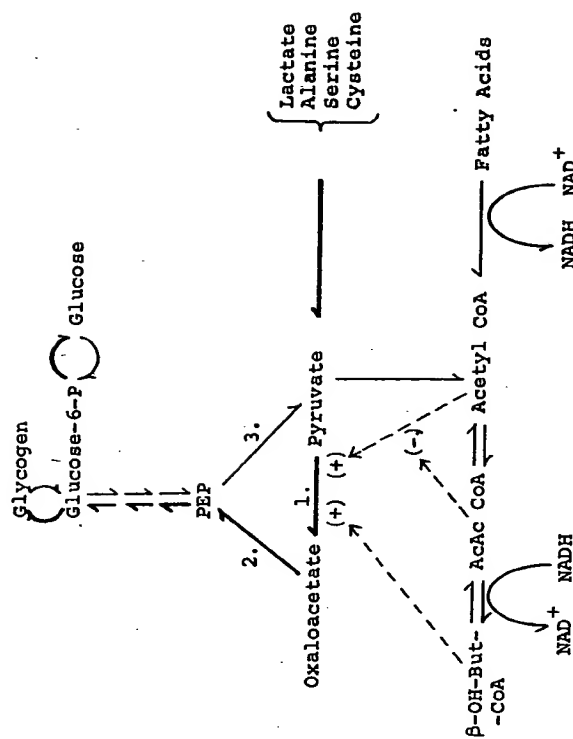
Fig. 20. Proposed role of pyruvate carboxylase in gluconeogenesis.

nutritional and physiological state (692–720). Furthermore, pyruvate carboxylase has been shown in vitro to be under the direct control of metabolite effectors, such as acetyl CoA, acetoacetyl CoA, β-hydroxybutyryl CoA, and ADP (108,641,642,644–646,672,721,722). Conditions, such as fasting (692–694), diabetes (893–700), and the administration of either adrenocorticosteroids (701–715) or glucagon (694,716–719), lead to both an increased rate of gluconeogenesis and an increase in pyruvate carboxylase level. Furthermore, adrenalectomy (712,714) or the administration of insulin to normal or diabetic animals (694,708, 720), which depresses gluconeogenesis, also causes a decrease in tissue carboxylase level. Dietary or physiological states associated with a high gluconeogenic rate and a high rate of mobilization and β-oxidation of long-chain fatty acids are characterized by elevated levels of acetyl CoA, an essential activator of pyruvate carboxylase (460,461,723–726). It has been demonstrated recently that acetoacetyl CoA concentrations decrease concurrently with increased gluconeogenesis mediated by glucagon (727). Since the K_A for acetyl CoA is increased by acetoacetyl CoA (672), amplification of acetyl CoA activation would be

achieved by lowered acetoacetyl CoA concentrations. This decrease in acetoacetyl CoA is correlated with an elevated mitochondrial redox state ($(NADH)/(NAD^+)$) (727) and would presumably be coupled to an increase in β-hydroxybutyryl CoA, a secondary activator of pyruvate carboxylase (672). Since these effectors appear to function antagonistically, more sensitive control over carboxylase activity is possible.

F. ATP:UREA AMIDOLYASE

Certain yeast and unicellular green algae, which lack a "classical urease," have the capacity to use urea as their sole nitrogen source (728–733). An increased biotin requirement for the utilization of urea by these yeast was observed (731,734–736), this requirement being reduced when other nitrogen sources were employed (736). Furthermore, the ability of whole cells (*Candida utilis*) to cleave urea was found to be correlated to their biotin content (733). It was demonstrated in Levenberg's laboratory (733,737–739) that cell-free extracts of *Candida utilis* and *Chlorella* catalyzed an avidin-sensitive ATP, Mg^{2+}-, and K^+-dependent cleavage of urea to CO_2, NH_3, ADP, and P_i. This biotin-dependent enzymatic activity could be induced by growth on urea and repressed by ammonia (733). A bicarbonate requirement for this reaction was demonstrated subsequently by Boon and Levenberg (109). The dependence of the enzymatic reaction on bicarbonate was not observed in initial studies because of its low K_m and its catalytic role in the reaction. The proposed 3-step partial reaction sequence catalyzed by ATP:urea amidolyase is shown below (109).

$$\text{Enz–biotin} + \text{HO—CO}_2^- + \text{ATP} \xrightarrow{Mg^{2+},K^+} \text{Enz–biotin–CO}_2^- + \text{ADP} + P_i \quad (28a)$$

$$\text{Enz–biotin–CO}_2^- + \text{H}_2\text{N—C(=O)—NH}_2 \rightleftharpoons \text{Enz–biotin} + \text{H}_2\text{N—C(=O)—NH—CO}_2^- \quad (28b)$$
$$\text{(urea)} \qquad \text{(N-carboxyurea (allophanate))}$$

$$\text{H}_2\text{N—C(=O)—NH—CO}_2^- + 2\text{HOH} \rightarrow 2\text{NH}_3 + \text{CO}_2 + \text{HO—CO}_2^- \quad (26c)$$

$$\text{Net:} \quad \text{ATP} + \text{H}_2\text{N—C(=O)—NH}_2 + 2\text{HOH} \xrightarrow{Mg^{2+},K^+} 2\text{NH}_3 + \text{CO}_2 + \text{ADP} + P_i \quad (28d)$$
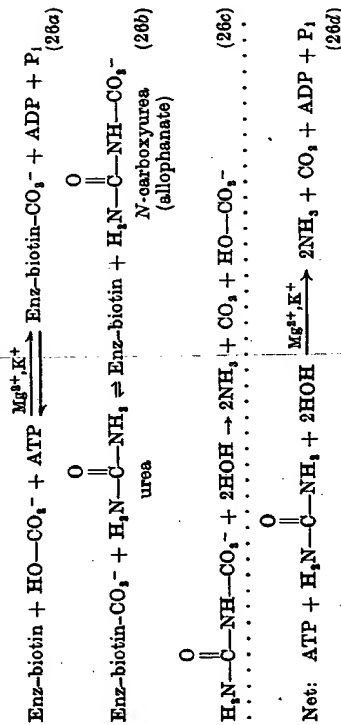
The first partial reaction (reaction 28a) is supported by the observation that the purified enzyme catalyzes a slow bicarbonate-dependent

cleavage of ATP to ADP and P_i, which is inhibited by avidin (109). The requirements for Mg^{2+} and K^+ appear to be restricted to the first partial reaction. Evidence for the second partial reaction (reaction 28b) is the finding that labilization of the ^{14}C-label in ^{14}C-urea to form $^{14}CO_2$ requires the addition of N-carboxyurea (allophanate). The enzyme exhibits a high degree of specificity toward allophanate as a carboxyl donor (109). The fact that this reaction occurs in the absence of the components of the first partial reaction indicates that the enzyme catalyzes an allophanate—^{14}C-urea exchange. The participation of biotin in this exchange reaction is shown by its sensitivity to avidin (109). The overall reversal of the first two steps (reactions 28a and b) was demonstrated by the allophanate-dependent formation of ATP from ADP and P_i (109). The terminal step (reaction 26c), unique among the reactions catalyzed by the biotin-dependent enzymes, is the avidin-insensitive hydrolytic decomposition of allophanate ($K_m = 6 \times 10^{-4}$ M) to NH_3 and CO_2 (109). In the overall forward reaction (reactions 26a,b,c), 2 moles of CO_2 are produced for every "CO_2" [presumably HCO_3^- (Section IV.A)] utilized. Consistent with this fact, Whitney and Cooper (740) observed that in the absence of added HCO_3^-, the overall reaction exhibited autocatalytic kinetics.

ATP:urea amidolyase has been purified from *Candida utilis* (109) and *Saccharomyces cerevisiae* (740). The enzyme from *Candida utilis* has been found to have a molecular weight of approximately 600,000 as determined by gel filtration (109), which is comparable to that of several other multisubunit biotin-dependent carboxylases (120,123, 547,567–570).

G. METHYLMALONYL CoA-PYRUVATE TRANSCARBOXYLASE

1. Background

Methylmalonyl CoA-pyruvate transcarboxylase catalyzes a unique terminal reaction in the fermentation of glucose, glycerol, or lactate to propionate by propionibacteria (602,604,741). In 1935 it was observed by Wood and Werkman (742) that net utilization of CO_2 occurred during the fermentation of glycerol to propionate carried out by this group of bacteria. CO_2, which is incorporated into the C-1 position of propionate (743), accounted for 50% of propionate carboxyl formed during the fermentation of glucose (744). Labeling experiments (744–750) indicated that a symmetrical intermediate, presumably

succinate, lies on the pathway between pyruvate and propionate. This proposal seemed reasonable since it had been demonstrated that succinate could be decarboxylated by these bacteria to give rise to propionate (751-754). However, the rate of propionate formation from succinate was found to be 10-50 times faster than that of CO_2 production (755), indicating that decarboxylation per se was not the principal process involved. This and other evidence (748,756-758) suggested that a C_1 unit other than CO_2 was being formed. It was demonstrated by Swick and Wood (113) that cell-free extracts of *Propionibacterium shermanii* catalyzed the isomerization of succinyl CoA to methylmalonyl CoA prior to carboxyl transfer from this product to pyruvate forming propionyl CoA and oxaloacetate. Hence the one-carbon unit derived from "succinate decarboxylation" was in fact conserved via transcarboxylation to pyruvate. The function of these reactions in the propionic acid fermentation of propionibacteria, which has been studied extensively in Wood's laboratory (741), is illustrated in Figure 21. As with propionyl CoA carboxylase (Section VI.B.1) (343,344,759) and methylmalonyl CoA decarboxylase (Section VI.H) (117), *S*-methylmalonyl CoA was shown to be the isomer utilized by the transcarboxylase (760-762). Similar to the pathway of propionate metabolism in animal tissues (Section VI.B.1), methylmalonyl CoA generated from succinyl CoA by a coenzyme B_{12}-dependent mutase (step 2, Fig. 21) has the R-configuration (761) and must undergo racemization (760) (step 2, Fig. 21) prior to transcarboxylation (step 3, Fig. 21) (763). The dependence of "succinate decarboxylation" on biotin first observed by Delwiche (29) with whole cells of *Propionibacterium pentosaceum* was later shown by Swick and Wood (113) to be due to the participation of biotin in the transcarboxylation reaction. The methylmalonyl CoA pyruvate transcarboxylase has been obtained in homogeneous form from *P. shermanii* and studied extensively in Wood's laboratory (763).

2. Reaction and Mechanism

The equilibrium constant

$$K_{eq} = \frac{[\text{propionyl CoA}][\text{oxaloacetate}]}{[S\text{-methylmalonyl CoA}][\text{pyruvate}]}$$

for the 2-step transcarboxylase-catalyzed reaction shown below is 0.53

Glucose → Pyruvate

[-CO₂]

Oxaloacetate — 4. (transcarboxylase) — S-Methylmalonyl CoA → Propionyl CoA

[H]

Malate

Fumarate — 3. (racemase) — R-Methylmalonyl CoA

[H]

Succinate — Succinyl CoA — 2. (mutase)

1. (CoA transferase) — [CoA] → Propionate
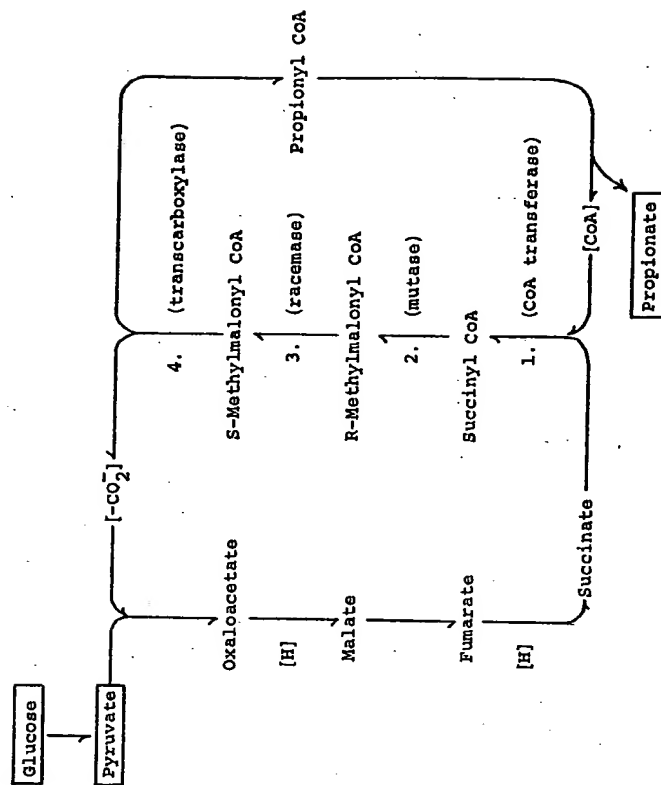
Fig. 21. Role of transcarboxylase in propionic acid fermentation of propionibacteria. From Allen et al. (741).

at pH 6.5 and 30°, which corresponds to a ΔF°_{303} of +0.39 kcal/mole (114).

Enz–biotin + (S)-methylmalonyl CoA ⇌ Enz–biotin–CO_2^- + propionyl CoA   (27a)

Enz–biotin–CO_2^- + pyruvate ⇌ Enz–biotin + oxaloacetate   (27b)

Net: (S)-methylmalonyl CoA + pyruvate ⇌ propionyl CoA + oxaloacetate   (27c)
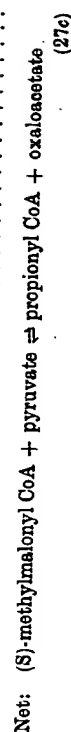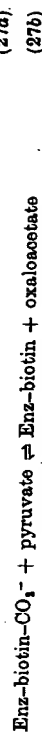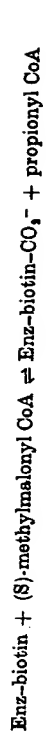
The enzyme has a broad specificity with respect to the acyl CoA substrates for the first partial reaction, hence malonyl CoA and ethylmalonyl CoA can serve as carboxyl donors in the forward direction and acetyl CoA and *n*-butyryl CoA as acceptors in the reverse direction (114,763). The relative rates for propionyl CoA, acetyl CoA, and *n*-butyryl CoA, 1, 0.55, and 0.093, respectively (114±), are attributable

to V_{max}, rather than K_m effects. In contrast, the enzyme is highly specific for keto acid substrates, α-ketobutyrate, α-ketovalerate, α-keto-glutarate, and β-ketoglutarate being inactive (114).

Support for the participation of a carboxylated enzyme intermediate in the overall reaction was obtained by Wood et al. (87) with the isolation of enzyme-$^{14}CO_2^-$ following a short incubation with methyl-malonyl-^{14}C-CoA. The exact equivalence between the biotin content of the enzyme and "labile-$^{14}CO_2$" suggested that the biotinyl prosthetic group was the site of carboxylation. The site of carboxylation was identified as the $1'$-N-position of the biotinyl moiety as described in Section IV.B.1 (87). Furthermore, it was demonstrated that the prosthetic group is linked through an amide bond to a lysyl ε-amino group of the apoprotein (Fig. 2; Section III.A and III.C) (87,181). As in the case of the carboxylated forms of the other biotin enzymes, spontaneous decarboxylation occurs (Table IV); the activation energy for this decarboxylation with transcarboxylase is 26.6 kcal (87). Furthermore, the $\Delta F°$ for the decarboxylation was found to be -4.74 kcal/mole (87), which indicates that the labile carboxyl group of enzyme-biotin-CO_2^- has a relatively high group transfer potential.

Further evidence for the occurrence of the half-reactions shown above (reactions 27a and b) is the transcarboxylase-catalyzed exchange between methylmalonyl CoA and ^{14}C-propionyl CoA which involves the first half-reaction (reaction 27a) and that between oxaloacetate and ^{14}C-pyruvate which involves the second partial reaction (reaction 27b). The exchange between the acyl CoA derivatives was observed in the absence of keto acids, and similarly the exchange between keto acids was independent of thioesters (91). These results are consistent with the independence of the two partial reactions as previously demonstrated for the acyl CoA and keto acid carboxylases (Section II.C.1. Analogous to the second step of the pyruvate carboxylase-catalyzed reaction (reaction 25b), which utilizes tightly bound metal for keto acid activation, Northrop and Wood (122) have found that the tightly bound metal of transcarboxylase serves a similar function in the comparable half-reaction (reaction 27b). Transcarboxylase contains both tightly bound cobalt and zinc (122), the paramagnetic properties of cobalt permitting the use of magnetic resonance techniques similar to those employed by Mildvan et al. (155) in studies with pyruvate carboxylase. Northrop and Wood (122) observed that enzyme-bound cobalt caused a broadening of the NMR signal of the

methyl protons of pyruvate and that oxalate, an inhibitor which binds tightly to the metal, reverses this effect. These results are consistent with the formation of an enzyme–metal–substrate bridge complex (122) as discussed earlier (Section IV.C.2) and illustrated in Figure 10 for pyruvate carboxylase (Section VI.E.2). It has been suggested (122) that the electron-withdrawing properties of the metal facilitate the departure of a methyl proton of pyruvate rendering it more susceptible to carboxylation. Since the catalytic activity of transcarboxylase does not vary greatly even though the proportions of zinc and cobalt differ markedly, it appears that cobalt and zinc may be functionally equivalent (122).

The results of the exchange studies discussed earlier which indicated that independent sites were involved in the two half-reactions are supported by kinetic studies (90). When pairs of substrates were compared in both the forward and reverse directions, the initial velocity patterns were in accord with the rate equation for a classical "ping-pong" mechanism (see mechanism 15 in Section V.A) (90). However, kinetic patterns with inhibitors of the overall reaction (90), as well as of the methylmalonyl CoA–^{14}C-propionyl CoA– and oxaloacetate–^{14}C-pyruvate exchange reactions (91) were not inter-pretable in terms of this kind of mechanism (90,91). Rather, these results are compatible with a hybrid "ping-pong" mechanism in which two distinct sites exist for the two sets of substrates but both sites can be occupied simultaneously. It has been proposed that the biotinyl prosthetic group can oscillate between these sites in the transcarboxyla-tion mechanism and that its occupancy of a particular site is influenced by substrate binding at the other site (91). Northrop and Wood (91) have demonstrated, for example, that 2-ketobutyrate, a pyruvate analogue, stimulates methylmalonyl CoA–^{14}C-propionyl CoA exchange, whereas oxalate, an analogue of oxaloacetate, inhibits this exchange (91). "Ping-pong" mechanisms have been reported for other biotin-dependent enzymes (92,93) and are supported by the finding that the *E. coli* acetyl CoA carboxylase can be resolved into subunits responsible for different partial reactions (95,96,100,100a) (Section V.A and B).

3. Molecular Characteristics

Methylmalonyl CoA-pyruvate transcarboxylase, which has been isolated in homogeneous form from cell-free extracts of *P. shermanii*, has a molecular weight of $670,000 \pm 40,000$ (762,764). The enzyme

contains 6 moles of biotin per mole of enzyme (762), and the sum of the tightly bound zinc and cobalt is 6 g atoms per mole (122). Although two active species of transcarboxylase, $s_{20,w} = \sim16\ S$ and $\sim18\ S$, are observed in sedimentation velocity experiments at rotor speeds (Spinco Model E ultracentrifuge) of 50,000 rpm or greater, at speeds of 30,000 rpm only one species, $s_{20,w} = \sim17\ S$, is observed (123). This phenomenon has been attributed to the pressure sensitivity of the transcarboxylase (123). Electron microscopic examination reveals that the native enzyme has a "Mickey Mouse" appearance, that is, a head with two large ears (765).

At low ionic strength and alkaline pH the enzyme dissociates into inactive 12-S and 6-S components (123,764); reconstitution of active transcarboxylase can be achieved at high ionic strength and neutral pH (123). Further dissociation of the 6-S component gives rise to a 1.3-S biotin-containing polypeptide and a component of undetermined size (123). Enzyme dissociated to this level can be reassociated to an enzymatically active transcarboxylase species which has an $s_{20,w}$ of $\sim25\ S$ (123) and appears in the electron microscope as a "Mickey Mouse," with extra ears (765).

The 1.3-S biotin-containing polypeptide has a molecular weight of approximately 11,700 by sedimentation equilibrium and a minimum molecular weight by amino acid analysis of 11,000 (121). This subunit contains 1 biotinyl prosthetic group and 1 sulfhydryl group (cysteinyl residue) per molecule with alanine as an amino terminus (121). This small biotin-containing polypeptide is similar to and probably serves the same function as the carboxyl-carrier protein of the *E. coli* acetyl CoA carboxylase system (Sections V.B, VI.A.2 and 3) (97).

H. METHYLMALONYL CoA DECARBOXYLASE

Methylmalonyl CoA decarboxylase, which catalyzes a terminal step in lactate fermentation of *Micrococcus lactilyticus* (766–770), was isolated from cell-free extracts of this organism by Galivan and Allen (116,771). In 1960 Phares and Carson (772) reported that a particulate fraction from *M. lactilyticus* decarboxylated methylmalonyl CoA to form propionyl CoA and CO_2; however, it was not clear which of three potential pathways was involved: (a) transcarboxylation yielding oxaloacetate and propionyl CoA, followed by oxaloacetate decarboxylation (Section VI.G), (b) reversal of the propionyl CoA carboxylase-catalyzed reaction (Section VI.B), or (c) direct decarboxylation of

methylmalonyl CoA (116). Galivan and Allen (116) showed that methylmalonyl CoA decarboxylase activity was sufficient to account for the rate of utilization (decarboxylation) of succinate by this organism. On the basis of this fact and other evidence (117) the following pathway was proposed:

$$\text{succinate} \xrightarrow{\text{(CoA transferase)}} \text{succinyl CoA} \xrightarrow{\text{(mutase)}} \text{(R)-methylmalonyl CoA} \xrightarrow{\text{(racemase)}} \text{(S)-methylmalonyl CoA} \xrightarrow{\text{(decarboxylase)}} \text{propionyl CoA} + CO_2 \rightarrow \boxed{\text{propionate}}$$
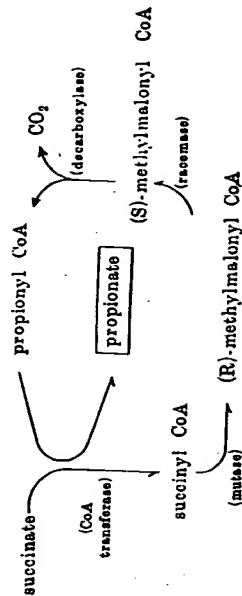
Although the decarboxylase was found to be tightly associated with the ribosomal fraction, it could be solubilized, but with considerable loss of activity. The possibility has been entertained that the ribosome-bound enzyme may be associated with the cell membrane (116), which would permit more rapid diffusion of fermentation products into the medium. The state of aggregation of the ribosome–enzyme complex, which was dependent on magnesium concentration, did not affect the activity of the enzyme. Furthermore, removal of RNA from the decarboxylase preparation did not result in decreased catalytic activity (117).

The soluble enzyme has an $s_{20,w}$ of 11.5 and an estimated molecular weight of 275,000–300,000 determined by the sucrose density gradient centrifugation method. The biotin content of this purified enzyme preparation was approximately 0.1 mole per 300,000 g of protein. This and the fact that the soluble and particulate preparations are avidin-sensitive (773) indicate that the decarboxylase is a biotin enzyme. Presumably, the enzyme functions via the intermediate formation of enzyme-biotin-CO_2^- (117):

Enz-biotin + (S)-methylmalonyl CoA $\rightleftharpoons$ Enz–biotin–CO_2^- + propionyl CoA   (28a)

Enz–biotin–CO_2^- + H$^+$ → Enz–biotin + CO_2 (28b)

Net: (S)-methylmalonyl CoA + H$^+$ → propionyl CoA + CO_2 (28c)
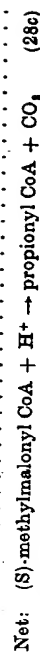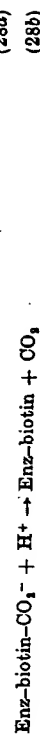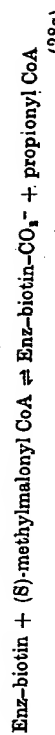
The observation that the enzyme catalyzes an avidin-sensitive methylmalonyl CoA-propionyl-1-^{14}C–CoA exchange supports the

Characteristics of the Biotin-dependent Enzymes

Biotin content	Metal content g atom/mole (native)	Catalytic center activity T.N./biotin[d]	Effector	References
1/4.1 × 10^4 g	0	3,000	Citrate (+)	119, 138, 401, 402, 406
0	—	1,100[e]	—	100
1/9 × 10^3 g	—	—	—	97
1/1.75 × 10^5 g	—	3,000	—	547
1/1.7 × 10^4 g	—	4,000	—	152, 548
1/1.9 × 10^4 g	—	2,000	—	138, 570
1/1.6 × 10^4 g	4 Mn	4,600	{Acetyl CoA (+), Acetoacetyl CoA (−), β-OH but. CoA (+)	120, 137, 157, 641, 642, 656-658, 667, 671, 672
1/1.5 × 10^6 g	4 Zn	3,600	{Aspartate (−), acetyl CoA (+), Palmit. CoA (+)	153, 336, 567-569, 628, 647-649, 656, 657, 674, 675, 678, 679
—	—	—	—	109
1/1.12 × 10^5 g	6 Co + Zn	5,700	—	122, 123, 762-765
1/1.2 × 10^3 g	—	—	—	121
—	—	—	—	117

Table V.

TABLE

Comparison of the Molecular and Catalytic

Enzyme	Molecular weight	$s_{20,w}$ (S)	Electron microscopic structure	Protomer Weight (daltons)	Protomer $s_{20,w}$ (S)	Minimal subunit Weight (daltons)	Minimal subunit $s_{20,w}$ (S)
Acetyl CoA carboxylase[a]							
Avian liver	4-10 × 10^6	59	Filaments	410,000	13.1	110,000	4.3
E. coli {BC[b]	100,000	5.7	—	—	—	50,000	—
{CCP[c]	9,000	1.3	—	—	—	—	—
Propionyl CoA carboxylase							
Pig heart	697,000	19.7	—	—	—	—	2.5
Bovine liver	—	19.0	—	—	—	—	—
β-Methylcrotonyl CoA Carboxylase	760,000	20.7	—	—	—	—	—
Pyruvate carboxylase							
Avian liver	660,000	14.8	Tetramer (square)	—	6.75	45,000	2.7
Yeast	600,000	15.8	Tetramer (diamond)	—	6-7	—	—
ATP:urea amidolyase	~800,000	—	—	—	—	—	—
Transcarboxylase	670,000	16.18	"Mickey mouse"	—	—	—	1.35
CCP[c]	12,000	1.3	—	—	—	—	—
Methylmalonyl CoA decarboxylase	275,000-300,000	11.5	—	—	—	—	—

[a] Properties of the avian liver and bovine adipose tissue carboxylases are compared in Table V.
[b] BC = biotin carboxylase component.
[c] CCP = carboxyl carrier protein component.
[d] T.N. = turnover number.
[e] Assuming 1 biotin binding site per 100,000.

occurrence of the first partial reaction (reaction 28a). Since racemic methylmalonyl CoA and enzymatically generated (S)-methylmalonyl CoA are decarboxylated 50% and 100%, respectively, it is apparent that only (S)-methylmalonyl CoA serves as substrate for the decarboxylase. It is interesting that propionyl CoA carboxylase (Section VI.B) and transcarboxylase (Section VI.G.) have the same stereochemical specificity. The decarboxylase appears to be specific for methylmalonyl CoA ($K_m = 10^{-6}\ M$), since malonyl CoA and succinyl CoA are not decarboxylated; however, both of these CoA derivatives are noncompetitive inhibitors having K_is of $10^{-4}\ M$ and $3 \times 10^{-4}\ M$, respectively.

The significance of the activation energy (10.8 kcal per mole) of the decarboxylase-catalyzed reaction in relation to that for the decarboxylation of enzyme-biotin-CO_2^- is discussed in Section IV.B.2 (Table IV).

I. OXALOACETATE DECARBOXYLASE

Stern (115) has isolated an inducible oxaloacetate decarboxylase from *Aerobacter aerogenes* which is tightly bound to the cytoplasmic membrane. Unlike the previously characterized oxaloacetate decarboxylases which exhibit a divalent cation requirement and are avidin-insensitive (115), the decarboxylase from *A. aerogenes* appears to be biotin-dependent and divalent cation-independent. Since the decarboxylation catalyzed by this enzyme does not require ADP or P_i, it is distinct from pyruvate carboxylase. Analogous to pyruvate carboxylase (Section VI.E) and transcarboxylase (Section VI.G.) which utilize keto acid substrates, oxaloacetate decarboxylase is probably a metalloenzyme. Certain metal binding agents such as cyanide and 8-hydroxyquinoline are inhibitory at relatively high concentration, although EDTA is not. Like all known biotin enzymes, the decarboxylase is almost completely inhibited by avidin and this inhibition is prevented by prior treatment of avidin with biotin. The probable role of biotin in the decarboxylation reaction is discussed in Section IV.B.2. Another interesting aspect of this enzyme is its relatively specific activation by Na+; Li+, NH4+, Rb+, Cs+, and K+ are inactive (115).

J. COMPARATIVE ASPECTS

The molecular and catalytic properties of the biotin-dependent enzymes are summarized in Table VII. All of these enzymes appear

to possess complex quaternary structures, being composed of nonidentical subunits. The biotinyl prosthetic group which functions as a "mobile carboxyl carrier" (Section V) is attached in amide linkage to a specific lysyl ε-amino group of the apoenzyme (Section III.C). In the case of two bacterial enzymes, transcarboxylase (Section VI.G.3) and acetyl CoA carboxylase (Section VI.A.3), this lysyl residue is part of a small carboxyl carrier peptide (molecular weight 12,000 and 9,000, respectively). In contrast, the carboxyl carrier peptide of the avian liver enzyme is a large structure of 100,000 daltons (Section VI.A.3). The catalytic center activities of the biotin-dependent enzymes are remarkably similar (Table VII) despite differences in assay conditions, for example, temperature or dependence on allosteric effectors.

The enzymes fall into two classes based on the nature of the substrate acted upon, that is, acyl CoA derivatives or keto acids. Those (pyruvate carboxylase and transcarboxylase) which utilize keto acid substrates contain tightly bound metal ion which functions in substrate activation (Section IV.C.2). There is a stoichiometric equivalence between the metal ion and biotin content of these metalloenzymes (Table VII).

VII. Other Proteins of Significance in the Biochemistry of Biotin

A. AVIDIN

Avidin, the specific biotin-binding protein from egg white, has served a useful function as a diagnostic tool in the identification and study of biotin-dependent enzymes. The interaction of avidin with biotin is unparalleled with regard to the strength of noncovalent binding between a protein and low molecular weight ligand, the dissociation constant of the binding sites being about $10^{-15}\ M$ (5). In addition to binding free biotin, this protein forms extremely tight complexes with the prosthetic group of biotin enzymes, leading to loss of catalytic activity. Avidin was identified as the heat-labile component in egg white (774–776) which was toxic to animals (2,777), causing a pellagra-like syndrome (778) and severe dermatitis (14). It is interesting that the avidin-biotin complex formed in the gastrointestinal tracts of animals is unavailable to the host (779,780), whereas the complex administered parenterally is apparently digested and biotin made available (781).

Avidin has been isolated from hen's egg white (782,783) and the oviduct of avian species (784) and amphibia (784), as well as from *Streptomyces* (streptavidin) (785,786). The synthesis of avidin, which occurs in the oviducts of laying hens, ceases with the cessation of egg-laying (787), and can be initiated by the administration of certain hormones, that is, deoxycorticosterone, testosterone, or stilbesterol (788) followed by progesterone (789,790).

1. Structure

In 1942 Woolley and Longworth (791) reported that avidin was a basic protein having an isoelectric point of about 10 and a molecular weight of 70,000 ($s_{20} = 4.7$ S). Early studies with highly active avidin preparations indicated that it was a glycoprotein (792–794); more recent studies have confirmed this observation (795–797). Some of the physical properties of avidin, later investigated in more detail in Green's laboratory (798,799) with homogeneous avidin preparations, are summarized in Table VIII. Avidin has a tetrameric structure as evidenced by its dissociation into weight-homogeneous subunits of about 17,000 daltons in the presence of 3–6 M guanidine·HCl (Table VIII). It is interesting that the avidin–(biotin)₄ complex is completely resistant to dissociation with 6 M guanidine·HCl. Investigations by DeLange (796) on the subunit structure indicate that the subunits are identical, have an alanine NH₂-terminus (800), and are composed of 129 amino acid residues, 4 mannose residues, and 3 glucosamine residues. Like most glycoproteins the carbohydrate moiety is attached to an asparaginyl residue (residue 17) (796). The earlier suggestion by Green (801) that there may be a genetic relationship between avidin, lysozyme, and α-lactalbumin does not appear likely in view of the amino acid sequence of avidin determined in DeLange's laboratory (795,796,796a, 796b,796c).

A tetrameric structure for avidin is consistent with the fact that each avidin molecule has 4 biotin-binding sites (799). Furthermore, studies by Green (802) with bifunctional biotin reagents, that is, bis-*N*-biotinyl-polymethylene diamines, reveal that the 4 biotin-binding sites are grouped in 2 pairs on opposite sides of the molecule (2:2 symmetry) which is consistent with X-ray crystallographic analysis of avidin crystals (803), as well as its appearance in the electron microscope (802). The minimal length of the chain linking the 2 biotinamide groups required for cross-linking 2 avidin molecules was C₁₀ to C₁₁

TABLE VIII
Physical Properties of Avidin[a]

	Protein concentration (mg/ml)	$s_{20,w}$ (S)	$\bar{v}$ (cm³/g)	$D_{20,w} \times 10^7$ (cm²/sec)	f/f_0	Molecular weight	Number of biotin-binding sites
Avidin	0.12– 8.3	4.55	0.72	5.98	1.33	67,900[b]	4[c]
Avidin–(biotin)₄ complex	7.4	4.58	—	6.08	—	—	—
Avidin + 6 M guanidine·HCl	8	0.76– 0.89	—	3.2– 4.5	3.7– 4.8	16,400– 19,500[b]	—
Avidin + 3 M guanidine·HCl	3	—	—	—	—	18,700[d]	—
Avidin–(biotin)₄ complex + 3–6 M guanidine·HCl	—	—	—	—	—	70,000[e]	—

[a] Data obtained from Green (798,799).
[b] M_{SD}
[c] $K_D = 10^{-15}$ M for avidin–(biotin)₄ complex.
[d] M_{SE}; similar molecular weights were obtained by the osmotic pressure method.
[e] Avidin–(biotin)₄ complex does not dissociate into subunits in the presence of guanidine·HCl as assessed by the spectrophotometric difference spectrum and fluorescence polarization (798).

(802). This suggests that the biotin is bound in clefts within or between subunits and that the bicyclic ring system of biotin is buried within this cleft 14–16 Å below the van der Waals surface of the avidin molecule.

2. Biotin-Binding Properties

As pointed out, biotin binds to avidin with great affinity (K_D, approximately 10^{-15} M, $k_{on} = 7 \times 10^7$/M-sec and $k_{off} = 9 \times 10^{-8}$/sec) and apparently at random among the 4 possible sites (5,804). Combination of biotin with avidin is accompanied by a free energy change of −20 kcal/mole of biotin bound with no net change in entropy; hence the reaction is exothermic ($\Delta H = -20.3$ kcal/mole) (805). It has been

suggested by Green (805) that the anticipated hydrophobic interactions for which there is some evidence (806,807) may be compensated for by a negative entropy change owing to the formation of buried hydrogen bonds. Importance has been attributed both to tryptophanyl residues [4 per subunit (797)] (804,806,808) and hydrogen bonds to the ureido ring (794) in the binding process. This is supported both by the fact that binding of biotin causes a red shift in the absorption spectrum which is characteristic of changes in the environment of tryptophan and because each biotin bound protects 4 tryptophan residues from oxidation by N-bromosuccinimide (NBS) (806,808,809). Destruction of 2 out of each 4 tryptophans with NBS essentially abolished biotin-binding ability (806). In support of a role for the ureido ring of biotin as a site for hydrogen bonding, it was found that modifying this ring system greatly weakened binding, whereas modifications of the tetrahydrothiophene ring or of the side-chain carboxyl group had little effect (806). On the other hand, the integrity of both the tetrahydro-thiophene and ureido rings appears to be crucial for biotin binding to strepavidin (810). Despite this difference, biotin binding to strepavidin results in changes in the absorption spectrum and tryptophan reactivity of strepavidin similar to those of avidin (811).

The binding of the biotinyl prosthetic group by avidin results in the complete loss of catalytic activity of all known biotin enzymes. However, the rate at which avidin combines with covalently bound biotin varies considerably depending on the accessibility of the prosthetic group. For example, the prosthetic group of the catalytically active form of avian liver acetyl CoA carboxylase in the presence of its allosteric activator, citrate, is virtually inaccessible to avidin presumably because in this state the biotinyl group is buried (139,347). On the other hand, the catalytically inactive form, that is, without effector, is rapidly inactivated (139). The reverse effect was noted with pyruvate carboxylase, that is, the presence of its allosteric activator, acetyl CoA, renders the prosthetic more accessible to avidin (671). Although it has been known for some time that the carboxylated forms of biotin enzymes (enzyme-biotin-CO_2^-) are inactivated by avidin (83), only recently was it realized that this is probably due to the fact that avidin promotes the decarboxylation of enzyme-biotin-CO_2^- (329). This effect has been interpreted in terms of the binding interaction causing strain in the carboxybiotinmolecule (Section IV.B.2).

3. Assay and Purification

The binding of biotin to avidin can be assessed by several methods based on (a) the ability of avidin to inhibit the growth of biotin-requiring organisms (813–818), (b) the binding of ¹⁴C-biotin (5,819), (c) the binding of an azo dye (4'-hydroxyazobenzene-2-carboxylic acid) in exact stoichiometry with biotin-binding sites followed by the spectrophotometric titration of its quantitative displacement from avidin with free or enzyme-bound biotin (807,812), and (d) the spectro-photometric titration of the biotin-induced red and blue shifts (tryptophan and tyrosine absorptions, respectively) of avidin (798,806, 812). The binding of ¹⁴C-biotin is the most sensitive of these procedures.

In addition to the standard methods for purifying avidin (796,820), Cuatrecasas and Wilchek (821) have developed a unique single-step purification procedure based on "affinity chromatography," which results in a 4000-fold purification and a 90% yield. Avidin added to a column of Biocytin-Sepharose,* is tightly bound to the column. After washing through most of the contaminating protein, the avidin is eluted in its dissociated form with 6 M guanidine·HCl at pH 1.5 and is then reconstituted by dilution. A modification of this technique has been used for the resolution of the biotin-containing subunit from the biotin-free subunits of liver acetyl CoA carboxylase (Section VI.A.3) (329).

B. BIOTINIDASE

In 1954 Thoma and Peterson (822) purified an enzyme from pig liver which catalyzed the release of covalently bound biotin from peptic digests of liver. Biotinidase (E.C. 3.5.1.12) has also been purified from *Streptococcus faecalis* (823), *Lactobacillus casei* (824), and pig kidney (824), the latter enzyme having been purified over 3000-fold. This enzyme is widely distributed in animal tissues; its presence in certain biotin-requiring bacteria enables them to utilize biocytin and other biotin derivatives for growth (823). Biotinidases are unable to act on native biotin-containing enzymes (823); however, following partial degradation, the biotinyl group becomes accessible to this enzyme. A number of biotin derivatives, including (+)-biotin amide,

* Prepared by coupling biocytin (ε-N-(+)-biotinyl-L-lysine) to cyanogen bromide-activated Sepharose (an agarose gel-filtration agent in bead form produced by Pharmacia).

(+)-biotin methyl ester, biocytin, N-(+)-biotinyl-p-aminobenzoate, N-(+)-biotinyl-β-alanine, (+)-biotinyl-L-aspartate, I'-N-methoxycarbonyl biocytin, and N-(— -biotinyl-p-aminobenzoate, are hydrolyzed by these enzymes (823,824). Like many other hydrolases biotinidase is inhibited by diisopropylfluorophosphate, suggesting the participation of a seryl hydroxyl or cysteinyl sulfhydryl group (824). The presence of hydroxylamine during hydrolysis leads to the formation of biotin hydroxamate (824). Biotinidase from hog kidney was used for the identification of the site of carboxylation of the biotinyl prosthetic group of several biotin enzymes (84–89).

Acknowledgments

We are indebted to Miss JoAnn Lopresti, Mrs. Carol Waters, and Mrs. Pat Lane for their skillful assistance in the preparation of the manuscript. We would like to acknowledge the cooperation of the many investigators who communicated their experimental findings to us prior to publication. We wish to express our gratitude to Pat, Martha, Carol, and Alan for their patience and perseverance during the writing of this paper.

References

1. Steinitz, F., Arch. Gesamte. Physiol., 72, 75 (1898).
2. Bateman, W. G., J. Biol. Chem., 26, 263 (1916).
3. Eakin, R. E., Snell, E. E., and Williams, R. J., J. Biol. Chem., 140, 535 (1941).
4. Launer, H. F., and Fraenkel-Conrat, H., J. Biol. Chem., 193, 125 (1951).
5. Green, N. M., Biochem. J., 89, 585 (1963).
6. Kögl, F., and Tönnis, B., Z. Physiol. Chem., 242, 43 (1936).
7. Kögl, F., Naturwissenschaften, 25, 465 (1937).
8. György, P., J. Biol. Chem., 131, 733 (1939).
9. György, P., Kuhn, R., and Lederer, E., J. Biol. Chem., 131, 745 (1939).
10. György, P., Melville, D. B., Burk, D., and du Vigneaud, V., Science, 91, 243 (1940).
11. du Vigneaud, V., Melville, D. B., György, P., and Rose, C. S., Science, 92, 62 (1940).
12. György, P., Rose, C. S., Hofmann, K., Melville, D. B., and du Vigneaud, V., Science, 92, 609 (1940).
13. Boas, M. A., Biochem. J., 21, 712 (1927).
14. Lease, J. G., and Parsons, H. T., Biochem. J., 28, 2109 (1934).

15. Allison, F. E., Hoover, S. R., and Burk, D., Science, 78, 217 (1933).
16. West, P. M., and Wilson, P. W., Science, 89, 607 (1939).
17. Robinson, F. A., The Vitamin B Complex, Wiley, New York, 1951, p. 404.
18. Shive, W., in R. J. Williams, R. E. Eakin, E. Beerstecher, Jr., and W. Shive, The Biochemistry of B Vitamins, Reinhold, New York, 1950, p. 542.
19. György, P., and Langer, B. W., Jr., in The Vitamins—Chemistry, Physiology, Pathology, Methods, Vol. II, W. H. Sebrell, Jr., and R. S. Harris, Eds., Academic Press, New York, 1968, p. 263.
20. Melville, D. B., Moyer, A. W., Hofmann, K., and du Vigneaud, V., J. Biol. Chem., 146, 487 (1942).
21. Harris, S. A., Wolf, D. E., Mozingo, R., and Folkers, K., Science, 97, 447 (1943).
22. Harris, S. A., Wolf, D. E., Mozingo, R., Anderson, R. C., Arth, G. E., Easton, N. R., Heyl, D., Wilson, A. N., and Folkers, K., J. Amer. Chem. Soc., 66, 1756 (1944).
23. Harris, S. A., Easton, N. R., Heyl, D., Wilson, A. N., and Folkers, K., J. Amer. Chem. Soc., 66, 1757 (1944).
24. Harris, S. A., Wolf, D. E., Mozingo, R., Arth, G. E., Anderson, R. C., Easton, N. R., and Folkers, K., J. Amer. Chem. Soc., 67, 2096 (1945).
25. Traub, W., Nature, 178, 649 (1956).
26. Bonnemere, C., Hamilton, J. A., Steinrauf, L. K., and Knappe, J., Biochemistry, 4, 240 (1965).
27. Trotter, J., and Hamilton, J. A., Biochemistry, 5, 713 (1966).
28. Lichstein, H. C., and Umbreit, W. W., J. Biol. Chem., 170, 329 (1947).
29. Delwiche, E. A., J. Bact., 59, 439 (1950).
30. Lichstein, H. C., J. Bact., 60, 485 (1950).
31. Lardy, H. A., Potter, R. L., and Elvehjem, C. A., J. Biol. Chem., 169, 451 (1947).
32. Potter, R. L., and Elvehjem, C. A., J. Biol. Chem., 172, 531 (1948).
33. Wessman, G. E., and Werkman, C. H., Arch. Biochem. Biophys., 26, 214 (1950).
34. Oser, B. L., Ann. Rev. Biochem., 17, 381 (1948).
35. Koser, S. A., Wright, M. H., and Dorfman, A., Proc. Soc. Exptl. Biol. Med., 51, 204 (1942).
36. Winzler, R. J., Burk, D., and du Vigneaud, V., Arch. Biochem., 5, 25 (1944).
37. Stokes, J. L., Larsen, A., and Gunness, M., J. Bact., 54, 219 (1947).
38. Stokes, J. L., Larsen, A., and Gunness, M., J. Biol. Chem., 167, 613 (1947).
39. Shive, W., and Rogers, L. L., J. Biol. Chem., 169, 453 (1947).
40. Stokstad, E. L. R., and Jukes, T. H., Ann. Rev. Biochem., 18, 452 (1949).
41. Lardy, H. A., Potter, R. L., and Burris, R. H., J. Biol. Chem., 179, 721 (1949).
42. White, J., and Munns, D. J., Nature, 165, 111 (1950).
43. Most, A. G., and Emmons, E. K., J. Bact., 68, 887 (1954).
44. Ahmad, F., and Rose, A. H., J. Gen. Microbiol., 28, 147 (1962).
45. Williams, V. R., and Fieger, E. A., J. Biol. Chem., 166, 335 (1946).
46. Hutchings, B. L., and Boggiano, E., J. Biol. Chem., 169, 229 (1947).

47. Axelrod, A. E., Hofmann, K., and Daubert, B. F., J. Biol. Chem., 169, 761 (1947).
48. Williams, V. R., and Fieger, E. A., J. Biol. Chem., 170, 399 (1947).
49. Williams, W. L., Broquist, H. P., and Snell, E. E., J. Biol. Chem., 170, 619 (1947).
50. Axelrod, A. E., Mitz, M., and Hofmann, K., J. Biol. Chem., 175, 265 (1948).
51. Cheng, A. L. S., Greenberg, S., Deuel, H. J., Jr., and Melnick, D., J. Biol. Chem., 192, 611 (1951).
52. Broquist, H. P., and Snell, E. E., Arch. Biochem. Biophys., 46, 432 (1953).
53. Mager, J., Kindler, S. H., and Grossowicz, N., J. Gen. Microbiol., 10, 130 (1954).
54. Lichstein, H. C., and Umbreit, W. W., J. Biol. Chem., 170, 423 (1947).
55. Lichstein, H. C., and Christman, J. F., J. Biol. Chem., 175, 649 (1948).
56. Lichstein, H. C., and Christman, J. F., J. Bact., 58, 565 (1949).
57. Ochoa, S., Mehler, A., Blanchard, M. L., Jukes, T. H., Hoffman, C. E., and Regan, M., J. Biol. Chem., 170, 413 (1947).
58. MacLeod, P. R., and Lardy, H. A., J. Biol. Chem., 179, 733 (1949).
59. MacLeod, P. R., Grisolia, S., Cohen, P. P., and Lardy, H. A., J. Biol. Chem., 180, 1003 (1949).
60. Blanchard, M. L., Korkes, S., del Campillo, A., and Ochoa, S., J. Biol. Chem., 187, 875 (1950).
61. Estes, J. M., Ravel, J. M., and Shive, W., J. Amer. Chem. Soc., 78, 6410 (1956).
62. Dixon, B., and Rose, A. H., Biochem. J., 99, 513 (1966).
63. Ahles, P. G., Ravel, J. M., and Shive, W., J. Biol. Chem., 236, 3263 (1961).
64. Langer, B. W., Jr., and Gyorgy, P., in The Vitamins—Chemistry, Physiology, Pathology, Methods, Vol. II, W. H. Sebrell, Jr., and R. S. Harris, Eds., Academic Press, New York, 1968, p. 335.
65. Welch, P. C., and Hartman, R. E., J. Bact., 100, 1138 (1969).
66. Lardy, H. A., and Peanasky, R., Physiol. Rev., 33, 560 (1953).
67. Titchener, E. B., Gibson, D. M., and Wakil, S. J., Federation Proc., 17, 322 (1958).
68. Wakil, S. J., Titchener, E. B., and Gibson, D. M., Biochim. Biophys. Acta, 29, 225 (1958).
69. Wakil, S. J., and Gibson, D. M., Biochim. Biophys. Acta, 41, 122 (1960).
70. Wright, L. D., and Skeggs, H. R., Proc. Soc. Exptl. Biol. Med., 56, 95 (1944).
71. Wright, L. D., Cresson, E. L., Skeggs, H. R., Wood, T. R., Peck, R. L., Wolf, D. E., and Folkers, K., J. Amer. Chem. Soc., 72, 1048 (1950).
72. Wright, L. D., Cresson, E. L., Skeggs, H. R., Peck, R. L., Wolf, D. E., Wood, T. R., Valiant, J., and Folkers, K., Science, 114, 635 (1951).
73. Wright, L. D., Cresson, E. L., Skeggs, H. R., Wood, T. R., Peck, R. L., Wolf, D. E., and Folkers, K., J. Amer. Chem. Soc., 74, 1996 (1952).
74. Peck, R. L., Wolf, D. E., and Folkers, K., J. Amer. Chem. Soc., 74, 1999 (1952).
75. Wolf, D. E., Valiant, J., Peck, R. L., and Folkers, K., J. Amer. Chem. Soc., 74, 2002 (1952).

76. Kosow, D. P., and Lane, M. D., Biochem. Biophys. Res. Commun., 5, 191 (1961).
77. Kosow, D. P., and Lane, M. D., Federation Proc., 21, 286 (1962).
78. Kosow, D. P., and Lane, M. D., Biochem. Biophys. Res. Commun., 7, 439 (1962).
79. Kosow, D. P., Huang, S. C., and Lane, M. D., J. Biol. Chem., 237, 3633 (1962).
80. Lynen, F., Knappe, J., Lorch, E., Jütting, G., and Ringelmann, E., Angew. Chem., 71, 481 (1959).
81. Lynen, F., Knappe, J., Lorch, E., Jütting, G., Ringelmann, E., and Lachance, J.P., Biochem. Z., 335, 123 (1961).
82. Knappe, J., Ringelmann, E., and Lynen, F., Biochem. Z., 335, 168 (1961).
83. Kaziro, Y., and Ochoa, S., J. Biol. Chem., 236, 3131 (1961).
84. Lane, M. D., and Lynen, F., Proc. Nat. Acad. Sci., U.S., 49, 379 (1963).
85. Knappe, J., Biederbick, K., and Brümmer, W., Angew. Chem., 74, 432 (1962).
86. Knappe, J., Wenger, B., and Wiegand, U., Biochem. Z., 337, 232 (1963).
87. Wood, H. G., Lochmüller, H., Riepertinger, C., and Lynen, F., Biochem. Z., 337, 247 (1963).
88. Numa, S., Ringelmann, E., and Lynen, F., Biochem. Z., 340, 228 (1964).
89. Numa, S., Bortz, W. M., and Lynen, F., Advan. Enzyme Regulation, 3, 407 (1965).
90. Northrop, D. B., J. Biol. Chem., 244, 5808 (1969).
91. Northrop, D. B., and Wood, H. G., J. Biol. Chem., 244, 5820 (1969).
92. Hashimoto, T., Iritani, N., Nakanishi, S., and Numa, S., Proc. Japanese Conf. on Biochem. Lipids Meetings, Ikaho, July 1970, p. 21.
93. Numa, S., Nakanishi, S., Hashimoto, T., Iritani, N., and Okazaki, T., Vitamins Hormones (1970), in press.
94. Alberts, A. W., and Vagelos, P. R., Federation Proc., 27, 647 (1968).
95. Alberts, A. W., and Vagelos, P. R., Proc. Nat. Acad. Sci. U.S., 59, 561 (1968).
96. Alberts, A. W., Nervi, A. M., and Vagelos, P. R., Proc. Nat. Acad. Sci. U.S., 63, 1319 (1969).
97. Nervi, A. M., and Alberts, A. W., Federation Proc., 29, 333 (1970).
98. Lane, M. D., Moss, J., Dimroth, P., Guchhait, R. B., Stoll, E., Edwards, J., and Kleinschmidt, A. K., Abstr. VIIIth Int. Congr. Biochem., Lucerne, Switzerland, p. 229 (1970).
99. Lane, M. D., Dimroth, P., Guchhait, R. B., and Stoll, E., Abstr. 160th Natl. Meeting Amer. Chem. Soc., Chicago, Biol. Chem. 98 (1970).
100. Dimroth, P., Guchhait, R. B., Stoll, E., and Lane, M. D., Proc. Nat. Acad. Sci. U.S., 67, 1353 (1970).
100a. Guchhait, R. B., Moss, J., Sokolski, W., and Lane, M. D., Proc. Nat. Acad. Sci. U.S., 68, 653 (1971).
100b. Dimroth, P., Guchhait, R. B., and Lane, M. D., Z. Physiol. Chem., 352, 351 (1971).
101. Brady, R. O., Proc. Nat. Acad. Sci., U.S., 44, 993 (1958).
102. Wakil, S. J., J. Amer. Chem. Soc., 80, 6465 (1958).

103. Formica, J. V., and Brady, R. O., J. Amer. Chem. Soc., 81, 752 (1959).
104. Flavin, M., Ortiz, P. J., and Ochoa, S., Nature, 176, 823 (1955).
105. Katz, J., and Chaikoff, I. L., J. Amer. Chem. Soc., 77, 2659 (1955).
106. Flavin, M., and Ochoa, S., J. Biol. Chem., 229, 965 (1957).
107. Lane, M. D., Halenz, D. R., Kosow, D. P., and Hegre, C. S., J. Biol. Chem., 235, 3082 (1960).
108. Utter, M. F., and Keech, D. B., J. Biol. Chem., 235, PC17 (1960).
109. Roon, R. J., and Levenberg, B., J. Biol. Chem., 245, 4593 (1970).
110. Lynen, F., Proc. Intern. Symp. Enzyme Chem., Maruzen, Tokyo, (1958), p. 57.
111. del Campillo-Campbell, A., Dekker, E. E., and Coon, M. J., Biochim. Biophys. Acta, 31, 290 (1959).
112. Seubert, W., Fass, E., and Remberger, U., Biochem. Z., 338, 265 (1963).
113. Swick, R. W., and Wood, H. G., Proc. Nat. Acad. Sci. U.S., 46, 28 (1960).
114. Wood, H. G., and Stjernholm, R., Proc. Nat. Acad. Sci. U.S., 47, 289 (1961).
115. Stern, J. R., Biochemistry, 6, 3545 (1967).
116. Galivan, J. H., and Allen, S. H. G., J. Biol. Chem., 243, 1253 (1968).
117. Galivan, J. H., and Allen, S. H. G., Arch. Biochem. Biophys., 126, 838 (1968).
118. Caplow, M., and Yager, M., J. Amer. Chem. Soc., 89, 4513 (1967).
119. Gregolin, C., Ryder, E., Warner, R. C., Kleinschmidt, A. K., Chang, H.-C., and Lane, M. D., J. Biol. Chem., 243, 4236 (1968).
120. Valentine, R. C., Wrigley, N. G., Scrutton, M. C., Irias, J. J., and Utter, M. F., Biochemistry, 5, 3111 (1966).
121. Gerwin, B. I., Jacobson, B. E., and Wood, H. G., Proc. Nat. Acad. Sci. U.S., 64, 1315 (1969).
122. Northrop, D. B., and Wood, H. G., J. Biol. Chem., 244, 5801 (1969).
123. Jacobson, B., Gerwin, B. I., Ahmad, F., Waegell, P., and Wood, H. G., J. Biol. Chem., 245, 6471 (1970).
124. Heinstein, P., and Stumpf, P. K., Proc. Australian Biochem. Soc., 5P (1969).
125. Heinstein, P. F., and Stumpf, P. K., J. Biol. Chem., 244, 5374 (1969).
126. Kaziro, Y., Leone, E., and Ochoa, S., Proc. Nat. Acad. Sci. U.S., 46, 1319 (1960).
127. Seubert, W., and Remberger, U., Biochem. Z., 334, 401 (1961).
128. Ochoa, S., and Kaziro, Y., Federation Proc., 20, 982 (1961).
129. Kaziro, Y., Hass, L. F., Boyer, P. D., and Ochoa, S., J. Biol. Chem., 237, 1460 (1962).
130. Scrutton, M. C., and Utter, M. F., Federation Proc., 23, 162 (1964).
131. Matsuhashi, M., Matsuhashi, S., and Lynen, F., Biochem. Z., 340, 263 (1964).
132. Scrutton, M. C., Keech, D. B., and Utter, M. F., J. Biol. Chem., 240, 574 (1965).
133. Ochoa, S., and Kaziro, Y., in Comprehensive Biochemistry, Vol. 16, M. Florkin and E. H. Stotz, Eds., American Elsevier, New York, 1965, p. 210.
134. Giorgio, A. J., and Plaut, G. W. E., Biochim. Biophys. Acta, 139, 487 (1967).
135. Lynen, F., Biochem. J., 102, 381 (1967).

136. Gregolin, C., Ryder, E., and Lane, M. D., J. Biol. Chem., 243, 4227 (1968).
137. Scrutton, M. C., and Utter, M. F., J. Biol. Chem., 240, 3714 (1965).
138. Hines, R. H., Young, D. L., Ringelmann, E., and Lynen, F., Biochem. Z., 337, 48 (1963).
139. Ryder, E., Gregolin, C., Chang, H.-C., and Lane, M. D., Proc. Nat. Acad. Sci. U.S., 57, 1455 (1967).
140. Stoll, E., Ryder, E., Edwards, J. B., and Lane, M. D., Proc. Nat. Acad. Sci. U.S., 60, 986 (1968).
141. Waite, M., and Wakil, S. J., Biochem. Biophys. Res. Commun., 9, 18 (1962).
142. Waite, M., and Wakil, S. J., J. Biol. Chem., 238, 77 (1963).
143. Waite, M., and Wakil, S. J., J. Biol. Chem., 238, 81 (1963).
144. Waite, M., and Wakil, S. J., J. Biol. Chem., 241, 1909 (1966).
145. Kaziro, Y., and Ochoa, S., Advan. Enzymol., 26, 283 (1964).
146. Knappe, J., Ann. Rev. Biochem., 39, 757 (1970).
147. Lane, M. D., and Halenz, D. R., Biochem. Biophys. Res. Commun., 2, 436 (1960).
148. Halenz, D. R., and Lane, M. D., Biochim. Biophys. Acta, 48, 426 (1961).
149. Friedman, D. L., and Stern, J. R., Biochem. Biophys. Res. Commun., 4, 266 (1961).
150. Halenz, D. R., and Lane, M. D., Biochem. Biophys. Res. Commun., 5, 27 (1961).
151. Hatch, M. D., and Stumpf, P. K., J. Biol. Chem., 236, 2879 (1961).
152. Halenz, D. R., Feng, J.-Y., Hegre, C. S., and Lane, M. D., J. Biol. Chem., 237, 2140 (1962).
153. Gailiusis, J., Rinne, R. W., and Benedict, C. R., Biochem. Biophys. Acta, 92, 595 (1964).
154. Matsuhashi, M., Matsuhashi, S., Numa, S., and Lynen, F., Biochem. Z., 340, 243 (1964).
155. Mildvan, A. S., Scrutton, M. C., and Utter, M. F., J. Biol. Chem., 241, 3488 (1966).
156. Scrutton, M. C., Utter, M. F., and Mildvan, A. S., Federation Proc., 24, 285 (1965).
157. Scrutton, M. C., Utter, M. F., and Mildvan, A. S., J. Biol. Chem., 241, 3480 (1966).
158. Northrop, D. B., and Wood, H. G., Federation Proc., 26, 343 (1967).
159. Hofmann, K., Melville, D. B., and du Vigneaud, V., J. Biol. Chem., 141, 207 (1941).
160. Brown, G. B., and du Vigneaud, V., J. Biol. Chem., 141, 85 (1941).
161. Melville, D. B., Hofmann, K., and du Vigneaud, V., Science, 94, 308 (1941).
162. du Vigneaud, V., Hofmann, K., Melville, D. B., and Rachele, J. R., J. Biol. Chem., 140, 763 (1941).
163. Kögl, F., and de Man, T. J., Z. Physiol. Chem., 269, 81 (1941).
164. Hofmann, K., Melville, D. B., and du Vigneaud, V., J. Amer. Chem. Soc., 63, 3237 (1941).
165. Kögl, F., and Pons, L., Z. Physiol. Chem., 269, 61 (1941).
166. du Vigneaud, V., Melville, D. B., Folkers, K., Wolf, D. E., Mozingo, R., Keresztesy, J. C., and Harris, S. A., J. Biol. Chem., 146, 475 (1942).

167. Hofmann, K., Kilmer, G. W., Melville, D. B., du Vigneaud, V., and Darby, H. H., J. Biol. Chem., 145, 503 (1942).
168. Kilmer, G. W., Armstrong, M. D., Brown, G. B., and du Vigneaud, V., J. Biol. Chem., 145, 495 (1942).
169. du Vigneaud, V., Hofmann, K., and Melville, D. B., J. Amer. Chem. Soc., 64, 188 (1942).
170. Melville, D. B., Hofmann, K., and du Vigneaud, V., J. Biol. Chem., 145, 101 (1942).
171. György, P., in The Vitamins, Vol. I, W. H. Sebrell, Jr., and R. S. Harris, Eds., Academic Press, New York, 1954, p. 532.
172. du Vigneaud, V., Science, 96, 455 (1942).
173. Hofmann, K., Melville, D. B., and du Vigneaud, V., J. Biol. Chem., 144, 513 (1942).
174. Harris, S. A., Mozingo, R., Wolf, D. E., Wilson, A. N., Arth, G. E., and Folkers, K., J. Amer. Chem. Soc., 66, 1800 (1944).
175. Traub, W., Science, 129, 210 (1959).
176. Shive, W., in The Biochemistry of B Vitamins, R. J. Williams, R. E. Eakin, E. Beerstecher, Jr., and W. Shive, Eds., Reinhold, New York, 1950, p. 544.
177. György, P., and Rose, C. S., Proc. Soc. Exptl. Biol. Chem., 49, 294 (1942).
178. Thompson, R. C., Eakin, R. E., and Williams, R. J., Science, 94, 589 (1944).
179. Bowden, J. W., and Peterson, W. H., J. Biol. Chem., 178, 533 (1949).
180. Hofmann, K., Dickel, D. F., and Axelrod, A. E., J. Biol. Chem., 183, 481 (1950).
181. Allen, S. H. G., Jacobson, B. E., and Stjernholm, R., Arch. Biochem. Biophys., 105, 494 (1964).
182. Kaziro, Y., Grossman, A., and Ochoa, S., J. Biol. Chem., 240, 64 (1965).
183. Mueller, J. H., J. Bact., 34, 163 (1937).
184. Mueller, J. H., J. Biol. Chem., 179, 121 (1937).
185. Mueller, J. H., Proc. Soc. Exptl. Biol. Med., 36, 706 (1937).
186. du Vigneaud, V., Dittmer, K., Hague, E., and Long, B., Science, 96, 186 (1942).
187. Eakin, R. E., and Eakin, E. A., Science, 96, 187 (1942).
188. Tatum, E. L., Ann. Rev. Biochem., 13, 667 (1944).
189. du Vigneaud, V., and Dittmer, K., cited by Duschinsky, R., and Dollan, L. A., J. Amer. Chem. Soc., 67, 2079 (1945).
190. Stokes, J. L., and Gunness, M., J. Biol. Chem., 157, 121 (1945).
191. Tatum, E. L., J. Biol. Chem., 160, 455 (1945).
192. Leonian, L. H., and Lilly, V. G., J. Bact., 49, 291 (1945).
193. Woolley, D. W., J. Biol. Chem., 183, 495 (1950).
194. Wright, L. D., and Driscoll, C. A., J. Amer. Chem. Soc., 76, 4999 (1954).
195. Wright, L. D., and Cresson, E. L., J. Amer. Chem. Soc., 76, 4156 (1954).
196. Wright, L. D., Cresson, E. L., and Driscoll, C. A., Proc. Soc. Exptl. Biol. Med., 89, 234 (1955).
197. Genghof, D. S., Arch. Biochem. Biophys., 62, 63 (1956).
198. Eisenberg, M. A., Federation Proc., 21, 467 (1962).
199. Eisenberg, M. A., J. Bact., 86, 673 (1963).
200. Rose, A. H., and Ilahi, M., Biochem. J., 89, 96P (1963).

201. Pai, C. H., and Lichstein, H. C., Biochim. Biophys. Acta, 100, 43 (1965).
202. Rose, A. H., Ilahi, M., and Kelemen, M. V., Biochem. J., 96, 319 (1965).
203. Eisenberg, M. A., and Maseda, R., Biochem. J., 101, 601 (1966).
204. Gray, C. H., and Tatum, E. L., Proc. Nat. Acad. Sci. U.S., 30, 404 (1944).
205. Schwartz, M., J. Bact., 92, 1083 (1966).
206. del Campillo-Campbell, A., Kayajanian, G., Campbell, A., and Adhya, S., J. Bact., 94, 2065 (1967).
207. Kayajanian, G., cited by ref. 206.
208. Kayajanian, G., Virology, 36, 30 (1968).
209. Rolfe, B., and Eisenberg, M. A., J. Bact., 96, 515 (1968).
210. Eisenberg, M. A., and Star, C., J. Bact., 96, 1291 (1968).
211. Nash, Ha., and Merrill, C. R., J. Mol. Biol., 43, 357 (1969).
212. Eisenberg, M. A., Biochem. Biophys. Res. Commun., 8, 437 (1962).
213. Elford, H. L., and Wright, L. D., Federation Proc., 21, 467 (1962).
214. Elford, H. L., and Wright, L. D., Biochem. Biophys. Res. Commun., 10, 373 (1963).
215. Eisenberg, M. A., Biochem. J., 98, 15C (1966).
216. Lezius, A., Ringelmann, E., and Lynen, F., Biochem. Z., 336, 510 (1963).
217. Eisenberg, M. A., Biochem. J., 101, 598 (1966).
218. Tepper, J. P., McCormick, D. B., and Wright, L. D., J. Biol. Chem., 241, 5734 (1966).
219. Iwahara, S., Kikuchi, M., Toohikura, T., and Ogata, K., Agr. Biol. Chem., 30, 304 (1966).
220. Eisenberg, M., and Maseda, R., Biochemistry, 9, 108 (1970).
221. Shemin, D., Russell, C. S., and Abramsky, T., J. Biol. Chem., 215, 613 (1955).
222. Eisenberg, M. A., and Krell, K., J. Biol. Chem., 244, 5503 (1969).
223. Eisenberg, M. A., and Krell, K., J. Bact., 98, 1227 (1969).
224. Pai, C. H., J. Bact., 99, 696 (1969).
225. Krell, K., and Eisenberg, M. A., Federation Proc., 28, 352 (1969).
226. Krell, K., and Eisenberg, M. A., J. Biol. Chem., 245, 6558 (1970).
227. Cooper, T. G., Filmer, D., Wishnick, M., and Lane, M. D., J. Biol. Chem., 244, 1081 (1968).
228. Pai, C. H., and Lichstein, H. C., J. Bact., 94, 1930 (1967).
229. Pai, C. H., and Lichstein, H. C., Arch. Biochem. Biophys., 114, 138 (1966).
230. Birnbaum, J., Pai, C. H., and Lichstein, H. C., J. Bact., 94, 1846 (1967).
231. Taylor, A. L., and Trotter, C. D., Bacteriol. Rev., 31, 332 (1967).
232. Rogers, T. O., and Lichstein, H. C., J. Bact., 100, 557 (1969).
233. Rogers, T. O., and Lichstein, H. C., J. Bact., 100, 565 (1969).
234. Pai, C. H., and Lichstein, H. C., Biochim. Biophys. Acta, 100, 28 (1965).
235. Pai, C. H., and Lichstein, H. C., Biochim. Biophys. Acta, 100, 36 (1965).
236. Lardy, H. A., and Adler, J., J. Biol. Chem., 219, 933 (1956).
237. Kosow, D. P., and Lane, M. D., Biochem. Biophys. Res. Commun., 4, 92 (1961).
238. Rominger, K. L., cited by ref. 260.
239. Jacobs, R., and Majerus, P. W., Federation Proc., 29, 735 (1970).
240. Gilgen, A., and Leuthardt, F., Helv. Chim. Acta, 45, 1833 (1962).

241. Foote, J. L., Christner, J. E., and Coon, M. J., *Federation Proc.*, 21, 239 (1962).

242. Mistry, S. P., and Dakshinamurti, K., *Biochem. Biophys. Acta*, 49, 606 (1961).

243. Wagle, S. R., *Arch. Biochem. Biophys.* 103, 297 (1963).

244. Suomalainen, H., and Keränen, A. J. A., *Suomen Kemistilehti B.*, 36, 88 (1963).

245. Lynen, F., and Rominger, K. L., *Federation Proc.*, 22, 537 (1963).

246. Foote, J. L., Christner, J. E., and Coon, M. J., *Biochim. Biophys. Acta*, 67, 676 (1963).

247. Schulthess, F. v., and Leuthard, T. F., *Helv. Chim. Acta*, 46, 1244 (1963).

248. Suomalainen, H., and Keränen, A. J. A., *Biochim. Biophys. Acta*, 70, 493 (1963).

249. Siegel, L., Foote, J. L., Christner, J. E., and Coon, M. J., *Biochem. Biophys. Res. Commun.*, 13, 307 (1963).

250. Vallotton, M., Borel, G. A., Schurter, R., and Leuthardt, F., *Helv. Chim. Acta*, 47, 2390 (1964).

251. Donaldson, W. E., *Proc. Soc. Exptl. Biol. Med.*, 116, 662 (1964).

252. Mistry, S. P., and Dakshinamurti, K., *Vitamins Hormones*, 22, 1 (1964).

253. Vollotton, M., and Leuthard, T. F., *Helv. Chim. Acta*, 47, 311 (1964).

254. Lane, M. D., Young, D. L., and Lynen, F., *J. Biol. Chem.*, 239, 2858 (1964).

255. Lane, M. D., Rominger, K. L., Young, D. L., and Lynen, F., *J. Biol. Chem.*, 239, 2865 (1964).

256. McAllister, H. C., Siegel, L., and Coon, M. J., *Federation Proc.*, 24, 540 (1965).

257. Siegel, L., Foote, J. L., and Coon, M. J., *J. Biol. Chem.*, 240, 1025 (1965).

258. Bhagavan, H. N., Coursin, D. B., and Dakshinamurti, K., *Arch. Biochem. Biophys.*, 110, 422 (1965).

259. Höpner, T., and Knappe, J., *Biochem. Z.*, 342, 190 (1965).

260. McAllister, H. C., and Coon, M. J., *J. Biol. Chem.*, 241, 2855 (1966).

261. Cazzulo, J. J., Sundaram, T. K., and Kornberg, H. L., *Nature*, 227, 1103 (1970).

262. Wagle, S. R., *Proc. Soc. Exptl. Biol. Med.*, 121, 15 (1966).

263. Varma, K. G., and Mistry, S. P., *Federation Proc.*, 26, 306 (1967).

264. Dakshinamurti, K., and Desjardins, P. R., *Can. J. Biochem.*, 46, 1281 (1968).

265. Dakshinamurti, K., and Desjardins, P. R., *Biochim. Biophys. Acta*, 176, 221 (1969).

266. Deodhar, A. D., and Mistry, S. P., *Biochem. Biophys. Res. Commun.*, 34, 755 (1969).

267. Birnbaum, J., *Arch. Biochem. Biophys.*, 132, 436 (1969).

268. Deodhar, A. D., and Mistry, S. P., *Arch. Biochem. Biophys.*, 131, 507 (1969).

269. Deodhar, A. D., and Mistry, S. P., *Arch. Biochem. Biophys.*, 129, 321 (1969).

270. Cazzulo, J. J., Sundaram, T. K., and Kornberg, H. L., *Nature*, 223, 1137 (1969).

271. Madappally, M. M., and Mistry, S. P., *Biochim. Biophys. Acta*, 215, 316 (1970).

272. Madappally, M. M., and Mistry, S. P., *Life Sciences*, 9, 833 (1970).

273. Mistry, S. P., and Madappally, M. M., *Federation Proc.*, 29, 294 (1970).

274. Arinze, J. C., and Mistry, S. P., *Comp. Biochem. Physiol.*, in press.

275. Nomoto, M., Narahashi, Y., and Murakami, M., *J. Biochem.*, 48, 593 (1960).

276. Nomoto, M., Narahashi, Y., and Murakami, M., *J. Biochem.*, 48, 906 (1960).

277. Wood, H. G., Lochmüller, H., and Lynen, F., *Federation Proc.*, 22, 537 (1963).

278. Reed, L. J., Leach, F. R., and Koike, M., *J. Biol. Chem.*, 232, 132 (1958).

279. Reed, L. J., in *The Enzymes*, Vol. 3, P. D. Boyer, H. Lardy, and K. Myrbäck, Eds., Academic Press, New York, 1960, p. 195.

280. Berg, P., *J. Amer. Chem. Soc.*, 77, 3163 (1955).

281. Peng, C. H. L., *Biochim. Biophys. Acta*, 22, 42 (1956).

282. Berg, P., *J. Biol. Chem.*, 222, 991 (1956).

283. Vignais, P. V., and Zabin, I., *Biochim. Biophys. Acta*, 29, 263 (1958).

284. Jencks, W. P., in *The Enzymes*, Vol. 6, P. D. Boyer, H. Lardy, and K. Myrbäck, Eds., Academic Press, New York, 1962, p. 373.

285. Webster, L. T., Jr., *J. Biol. Chem.*, 238, 4010 (1963).

286. Webster, L. T., Jr., and Campagnari, F., *Federation Proc.*, 22, 536 (1963).

287. Krishnaswamy, P. R., and Meister, A., *J. Biol. Chem.*, 235, 408 (1960).

288. Wong, K. K., and Moldave, K., *J. Biol. Chem.*, 235, 694 (1960).

289. Stulberg, M. P., and Novelli, G. D., in *The Enzymes*, Vol. 6, P. D. Boyer, H. Lardy, and K. Myrbäck, Eds., Academic Press, New York, 1962, p. 401.

290. Christner, J. E., Schlesinger, M. J., and Coon, M. J., *J. Biol. Chem.*, 239, 3997 (1964).

291. Khorana, H. G., *Some Recent Developments in the Chemistry of Phosphate Esters of Biological Interest*, Wiley, New York, 1961, p. 79.

292. Maruyama, H., and Lane, M. D., *Biochem. Biophys. Res. Commun.*, 9, 461 (1962).

293. Maruyama, H., Easterday, R. L., Chang, H.-C., and Lane, M. D., *J. Biol. Chem.*, 241, 2405 (1966).

294. Lane, M. D., in *CO₂: Chemical, Biochemical and Physiological Aspects*, R. E. Forster, J. T. Edsall, A. B. Otis, and F. J. W. Roughton, Eds., NASA, SP-188, U.S. Government Printing Office, Washington, D.C., 1969, p. 195.

295. Cooper, T. G., Tchen, T. T., Wood, H. G., Benedict, C. R., and Filmer, D. L., in *CO₂: Chemical, Biochemical, and Physiological Aspects*, R. E. Forster, J. T. Edsall, A. B. Otis, and F. J. W. Roughton, Eds., NASA, SP-188, U.S. Government Printing Office, Washington, D.C., 1969, p. 183.

295a. Filmer, D., and Cooper, T. G., *J. Theoretical Biol.*, 29, 131 (1970).

296. Cooper, T. G., Tchen, T. T., Wood, H. G., and Benedict, C. R., *J. Biol. Chem.*, 243, 3857 (1968).

297. Cooper, T. G., Tchen, T. T., Benedict, C. R., and Wood, H. G., *Federation Proc.*, 27, 587 (1968).

298. Scrutton, M. C., and Lane, M. D., in *CO₂: Chemical, Biochemical and Physiological Aspects*, R. E. Forster, J. T. Edsall, A. B. Otis, and F. J. W. Roughton, Eds., NASA, SP-188, U.S. Government Printing Office, Washington, D.C., 1969, p. 222.

299. Edsall, J. T., *Harvey Lec. Ser.*, 62, 191 (1966–1967).

300. Edsall, J. T., in CO_2: Chemical, Biochemical, and Physiological Aspects, R. E. Forster, J. T. Edsall, A.B. Otis, and F.J. W. Roughton, Eds, NASA, SP-188, U.S. Government Printing Office, Washington, D.C., 1969, p. 21.
301. Stiles, M., Ann. N.Y. Acad. Sci., 88, 332 (1960).
302. Stiles, M., J. Amer. Chem. Soc., 81, 2598 (1959).
303. Stiles, M., and Finkbeiner, H. L., J. Amer. Chem. Soc., 81, 505 (1959).
304. Finkbeiner, H. L., and Stiles, M., J. Amer. Chem. Soc., 85, 616 (1963).
305. Finkbeiner, H. L., J. Amer. Chem. Soc., 87, 4588 (1965).
306. Pelletier, S. W., Chappell, R. L., Parthasarathy, P. C., and Lewin, N., J. Org. Chem., 31, 1747 (1966).
307. Eigen, M., Kustin, K., and Maass, G., Z. Physik. Chem. (N.F.), 30, 130 (1961).
308. Krebs, H. A., and Roughton, R. J. W., Biochem. J., 43, 550 (1948).
309. Ryder, E., and Lane, M. D., unpublished observations.
310. Griffith, D. L., and Stiles, M., J. Amer. Chem. Soc., 87, 3710 (1965).
311. Lane, M. D., Edwards, J., Stoll, E., and Moss, J., Vitamins Hormones (1970), in press.
312. Olah, G. A., and White, A. M., J. Amer. Chem. Soc., 90, 6087 (1968).
313. Sigel, H., McCormick, D. B., Griesser, R., Prijs, B., and Wright, L. D., Biochemistry, 8, 2687 (1969).
314. Duschinsky, R., Dolan, L. A., Flower, D., and Rubin, S. H., Arch. Biochem., 6, 480 (1945).
315. Hofmann, K., McCoy, R. H., Felton, J., Axelrod, A. E., and Pilgrim, F. J., Arch. Biochem., 7, 393 (1945).
316. Pilgrim, F. J., Axelrod, A. E., Winnick, T., and Hofmann, K., Science, 102, 35 (1945).
317. McCoy, R. H., Felton, J. R., and Hofmann, K., Arch. Biochem. Biophys., 9, 141 (1946).
318. Hofmann, K., and Winnick, T., J. Biol. Chem., 160, 449 (1945).
319. Winnick, T., Hofmann, K., Pilgrim, F. J., and Axelrod, A. E., J. Biol. Chem., 161, 405 (1945).
320. Hofmann, K., and Axelrod, A. E., Arch. Biochem., 11, 375 (1946).
321. Axelrod, A. E., Flinn, B. C., and Hofmann, K., J. Biol. Chem., 169, 195 (1947).
322. Krueger, K. K., and Peterson, W. H., J. Bact., 55, 693 (1948).
323. Caplow, M., J. Amer. Chem. Soc., 87, 5774 (1965).
324. Bruice, T. C., and Hegarty, A. F., Proc. Nat. Acad. Sci. U.S., 65, 805 (1970).
324a. Guchhait, R. B., Sokolski, W., and Lane, M. D., unpublished observations (1971).
325. Knappe, J., Abstr. VIth Intern. Congr. Biochem., New York, 1964, p. 355.
326. Hammond, G. S., J. Amer. Chem. Soc., 77, 334 (1955).
327. Ryder, E., Gregolin, C., and Lane, M. D., unpublished observations.
328. Jencks, W. P., Catalysis in Chemistry and Enzymology, McGraw-Hill, New York, 1969, p. 286.
329. Edwards, J. B., and Lane, M. D., unpublished observations.
330. Schaeffer, H. J., and Bhargava, P. S., J. Pharm. Sci., 53, 137 (1964).
331. Knappe, J., and Lynen, F., Colloq. Ges. Physiol. Chem., 14th, 265 (1963).

332. Caplow, M., Biochemistry, 8, 2656 (1969).
333. Bowen, C. E., Rauscher, E., and Ingraham, L. L., Arch. Biochem. Biophys., 125, 865 (1968).
334. Bowen, C. E., and Ingraham, L. L., personal communication, manuscript in preparation.
335. Glasel, J. A., Biochemistry, 5, 1851 (1966).
336. Scrutton, M. C., and Young, M. R., Federation Proc., 29, 597 (1970).
337. Mildvan, A. S., and Scrutton, M. C., Biochemistry, 6, 2978 (1967).
338. Mildvan, A. S., Gregolin, C., and Lane, M. D., unpublished observations.
339. Prescott, D. J., and Rabinowitz, J. L., J. Biol. Chem., 243, 1551 (1968).
340. Prescott, D. J., and Rabinowitz, J. L., Federation Proc., 26, 562 (1967).
341. Gregolin, C., and Lane, M. D., unpublished observations.
342. Hegre, C. S., doctoral thesis, submitted to the Graduate Faculty, Virginia Polytechnic Institute, 1963.
343. Sprecher, M., Clark, M. J., and Sprinson, D. B., Biochem. Biophys. Res. Commun., 15, 581 (1964).
344. Rétey, J., and Lynen, F., Biochem. Biophys. Res. Commun., 16, 358 (1964).
345. Rétey, J., and Lynen, F., Biochem. Z., 342, 256 (1965).
346. Arigoni, D., Lynen, F., and Rétey, J., Helv. Chim. Acta, 49, 311 (1966).
347. Moss, J., and Lane, M. D., manuscript in preparation.
348. Eggerer, H., Biochem. Z., 343, 111 (1965).
349. Eggerer, H., and Klette, A., European J. Biochem., 1, 447 (1967).
350. Bové, J., Martin, R. O., Ingraham, L. L., and Stumpf, P. K., J. Biol. Chem., 234, 999 (1959).
351. Marcus, A., and Vennesland, B., J. Biol. Chem., 233, 727 (1958).
352. Bridger, W. A., Millen, W. A., and Boyer, P. D., Biochemistry, 7, 3608 (1968).
353. Lynen, F., Federation Proc., 12, 683 (1953).
354. Bordwell, F. G., and Boutan, P. J., J. Amer. Chem. Soc., 78, 854 (1956).
355. Bordwell, F. G., and Cooper, G. D., J. Amer. Chem. Soc., 74, 1058 (1952).
356. Baker, A. W., and Shulgrin, A. T., cited by ref. 350.
357. Talalay, P., and Wang, V. S., Biochim. Biophys. Acta, 18, 300 (1955).
358. Rose, I. A., and O'Connell, E. L., J. Biol. Chem., 236, 3086 (1961).
359. Rose, I. A., J. Biol. Chem., 245, 6052 (1970).
360. Cleland, W. W., Biochem. Biophys. Acta, 67, 104 (1963).
361. Cleland, W. W., Biochem. Biophys. Acta, 67, 173 (1963).
362. Cleland, W. W., Biochem. Biophys. Acta, 67, 188 (1963).
363. Wakil, S. J., Porter, J. W., and Gibson, D. M., Biochim. Biophys. Acta, 24, 453 (1957).
364. Porter, J. W., Wakil, S. J., Tietz, A., Jacob, M. I., and Gibson, D. M., Biochim. Biophys. Acta, 25, 35 (1957).
365. Porter, J. W., and Tietz, A., Biochim. Biophys. Acta, 25, 41 (1957).
366. Tietz, A., Biochim. Biophys. Acta, 25, 303 (1957).
367. Titchener, E. B., and Gibson, D. M., Federation Proc., 16, 262 (1957).
368. Gibson, D. M., Titchener, E. B., and Wakil, S. J., Biochim. Biophys. Acta, 30, 376 (1958).

369. Gibson, D. M., Titchener, E. B., and Wakil, S. J., *J. Amer. Chem. Soc.*, *80*, 2908 (1958).

370. Wakil, S. J., Titchener, E. B., and Gibson, D. M., *Biochim. Biophys. Acta*, *34*, 227 (1959).

371. Wakil, S. J., *J. Lipid Res.*, *2*, 1 (1961).

372. Wakil, S. J., *Vth Int. Congr. Biochem. Moscow, Symposium VII*, 1 (1961).

373. Fritz, I. B., *Physiol. Rev.*, *41*, 52 (1961).

374. Lynen, F., *Federation Proc.*, *20*, 941 (1961).

375. Vagelos, P. R., *Ann. Rev. Biochem.*, *33*, 139 (1964).

376. Wakil, S. J., and Ganguly, J., *J. Amer. Chem. Soc.*, *81*, 2597 (1959).

377. Ganguly, J., *Biochim. Biophys. Acta*, *40*, 110 (1960).

378. Squires, C. L., Stumpf, P. K., and Schmid, C., *Plant Physiol.*, *33*, 365 (1958).

379. Lynen, F., *J. Cell. Comp. Physiol.*, *54* (Suppl. 1), 33 (1959).

380. Brady, R. O., and Gurin, S., *J. Biol. Chem.*, *199*, 421 (1952).

381. Popják, G., and Tietz, A., *Biochem. J.*, *60*, 147 (1955).

382. Brady, R. O., Mamoon, A.-M., and Stadtman, E. R., *J. Biol. Chem.*, *222*, 795 (1956).

383. Dituri, F., Shaw, W. N., Warms, J. V. B., and Gurin, S., *J. Biol. Chem.*, *226*, 407 (1957).

384. Langdon, R. G., *J. Biol. Chem.*, *226*, 615 (1957).

385. Abraham, S., Matthes, K. J., and Chaikoff, I. L., *J. Biol. Chem.*, *235*, 2551 (1960).

386. Abraham, S., Matthes, K. J., and Chaikoff, I. L., *Biochim. Biophys. Acta*, *49*, 268 (1961).

387. Dils, R., and Popják, G., *Biochem. J.*, *80*, 47p (1961).

388. Martin, D. B., and Vagelos, P. R., *J. Biol. Chem.*, *237*, 1787 (1962).

389. Dils, R., and Popják, G., *Biochem. J.*, *83*, 41 (1962).

390. Lane, M. D., and Moss, J., in *Metabolic Pathways*, *5*, 23 (1971).

391. Wakil, S. J., and Ganguly, J., *Federation Proc.*, *18*, 346 (1959).

392. Waite, M., and Wakil, S. J., *J. Biol. Chem.*, *237*, 2750 (1962).

393. Martin, D. B., and Vagelos, P. R., *Federation Proc.*, *21*, 289 (1962).

394. Martin, D. B., and Vagelos, P. R., *Biochem. Biophys. Res. Commun.*, *7*, 101 (1962).

395. Waite, M., *Federation Proc.*, *21*, 287 (1962).

396. Kallen, R. B., and Lowenstein, J. M., *Arch. Biochem. Biophys.*, *96*, 188 (1962).

397. Matsuhashi, M., Matsuhashi, S., Numa, S. and Lynen, F., *Federation Proc.*, *21*, 288 (1962).

398. Abraham, S., Lorch, E., and Chaikoff, I. L., *Biochem. Biophys. Res. Commun.*, *7*, 190 (1960).

399. Vagelos, P. R., Alberts, A. W., and Martin, D. B., *Biochem. Biophys. Res. Commun.*, *8*, 4 (1962).

400. Vagelos, P. R., Alberts, A. W., and Martin, D. B., *J. Biol. Chem.*, *238*, 533 (1963).

401. Gregolin, C., Ryder, E., Kleinschmidt, A. K., Warner, R. C., and Lane, M. D., *Proc. Nat. Acad. Sci. U.S.*, *56*, 148 (1966).

402. Gregolin, C., Ryder, E., Warner, R. C., Kleinschmidt, A. K., and Lane, M. D., *Proc. Nat. Acad. Sci. U.S.*, *56*, 1751 (1966).

403. Numa, S., Ringelmann, E., and Riedel, B., *Biochem. Biophys. Res. Commun.*, *24*, 750 (1966).

404. Goto, T., Ringelmann, E., Riedel, B., and Numa, S., *Life Sciences*, *6*, 785 (1967).

405. Numa, S., Goto, T., Ringelmann, E., and Riedel, B., *European J. Biochem.*, *3*, 124 (1967).

406. Kleinschmidt, A. K., Moss, J., and Lane, M. D., *Science*, *166*, 1276 (1969).

407. Moss, J., Yamagishi, M., Kleinschmidt, A. K., and Lane, M. D., *Federation Proc.*, *28*, 1548 (1969).

408. Miller, A. L., and Levy, H. R., *J. Biol. Chem.*, *244*, 2334 (1969).

409. Shrago, E., Spennetta, T., and Gordon, E., *J. Biol. Chem.*, *244*, 2761 (1969).

410. White, D., and Klein, H. P., *Biochem. Biophys. Res. Commun.*, *20*, 78 (1965).

411. White, D., and Klein, H. P., *J. Bact.*, *91*, 1218 (1966).

412. Rasmussen, R. K., and Klein, H. P., *Biochem. Biophys. Res. Commun.*, *28*, 415 (1967).

413. Burton, D., and Stumpf, P. K., *Arch. Biochem. Biophys.*, *117*, 604 (1966).

414. Heinstein, P. F., and Stumpf, P. K., *Federation Proc.*, *27*, 647 (1968).

415. Gregolin, C., Stoll, E., and Lane, M. D., manuscript in preparation.

416. Moss, J., Yamagishi, M., and Lane, M. D., manuscript in preparation.

417. Gregolin, C., Ryder, E., Kleinschmidt, A. K., Warner, R. C., and Lane, M. D., *Abstr. 152nd Meeting Amer. Chem. Soc., New York*, C52 (1966).

418. Kleinschmidt, A. K., and Lane, M. D., *Biophys. J.*, *9*, A106 (1969).

419. Numa, S., personal communication (1970).

420. Bloom, W., and Fawcett, D. W., *A Textbook of Histology*, ninth edition, Saunders, Philadelphia, 1968, p. 169.

421. Schotz, M. C., Stewart, J. E., Garfinkel, A. S., Whelan, C. F., and Baker, N., in *Drugs Affecting Lipid Metabolism*, W. L. Holmes, L. A. Carlson, and R. Paoletti, Eds., Plenum, New York, 1969, p. 161.

422. Sheldon, H., personal communication (1969).

423. Lane, M. D., Gregolin, C., Ryder, E., Kleinschmidt, A. K., and Warner, R. C., *Abstr. 154th Meeting Amer. Chem. Soc., Chicago*, C261 (1967).

424. Ryder, E., Gregolin, C., Chang, H.-C., Kleinschmidt, A. K., Warner, R. C., and Lane, M. D., *Federation Proc.*, *26*, 672 (1967).

425. Lane, M. D., Ryder, E., Gregolin, C., Chang, H.-C., Kleinschmidt, A. K., and Warner, R. C., *Abstr. VII Int. Congr. Biochem., Tokyo*, (1967).

426. Guchhait, R. B., Dimroth, P., and Lane, M. D., manuscript in preparation.

427. Vagelos, P. R., personal communication (1970).

428. Birch, A. J., and Donovan, F. W., *Australian J. Chem.*, *6*, 360 (1953).

429. Birch, A. J., Massy-Westropp, R. A., and Moye, C. J., *Australian J. Chem.*, *8*, 539 (1955).

430. Dimroth, P., Walter, H., and Lynen, F., *European J. Biochem.*, *13*, 98 (1970).

431. Chang, H.-C., Seidman, I., Teebor, G., and Lane, M. D., *Biochem. Biophys. Res. Commun.*, *28*, 682 (1967).

432. Numa, S., Matsuhashi, M., and Lynen, F., *Biochem. Z.*, *334*, 203 (1961).
433. Kornacker, M. S., and Lowenstein, J. M., *Science*, *144*, 1027 (1964).
434. Lowenstein, J. M., in *Control of Energy Metabolism, A Colloquium of the Johnson Research Foundation*, B. Chance, R. W. Estabrook, and J. R. Williamson, Eds., Academic Press, New York, 1965, p. 261.
435. Spencer, A. F., and Lowenstein, J. M., *Biochem. J.*, *99*, 760 (1966).
436. Frohman, C. E., Orten, J. M., and Smith, A. H., *J. Biol. Chem.*, *193*, 803 (1951).
437. Start, C., and Newsholme, E. A., *Biochem. J.*, *107*, 411 (1968).
438. Ballard, F. J., Hanson, R. W., Kronfeld, D. S., and Raggi, F., *J. Nutr.*, *95*, 160 (1960).
439. Herrera, E., and Freinkel, N., *Biochim. Biophys. Acta*, *170*, 244 (1968).
440. Start, C., and Newsholme, E. A., *Biochem. J.*, *104*, 46P (1967).
441. Spencer, A. F., and Lowenstein, J. M., *Biochem. J.*, *103*, 342 (1967).
442. Parmeggiani, A., and Bowman, R. H., *Biochem. Biophys. Res. Commun.*, *12*, 268 (1963).
443. Dixit, P. K., DeVilliers, D. C., Jr., and Lazarow, A., *Metabolism*, *16*, 285 (1967).
444. Williamson, J. R., Herczeg, B., Coles, H., and Danish, R., *Biochem. Biophys. Res. Commun.*, *24*, 437 (1966).
445. Lynen, F., Matsuhashi, M., Numa, S., and Schweizer, E., in *The Control of Lipid Metabolism, Biochem. Soc. Symp.* 24, J. K. Grant, Ed., Academic Press, New York, 1963, p. 43.
446. Medes, G., Thomas, A., and Weinhouse, S., *J. Biol. Chem.*, *197*, 181 (1952).
447. Lyon, I., Masri, M. S., and Chaikoff, I. L., *J. Biol. Chem.*, *196*, 25 (1952).
448. Masoro, E. J., Chaikoff, I. L., Chernick, S. S., and Felts, J. M., *J. Biol. Chem.*, *185*, 845 (1950).
449. Guchhait, R. B., and Mukherjee, S., manuscript in preparation.
450. Wieland, O., Weiss, L., and Eger-Neufeldt, I., *Abstr. VIth Int. Congr. Biochem.*, New York, *9*, 709 (1964).
451. Muto, Y., and Gibson, D. M., *Biochem. Biophys. Res. Commun.*, *38*, 9 (1970).
452. Hill, R., Linazasoro, J. M., Chevallier, F., and Chaikoff, I. L., *J. Biol. Chem.*, *233*, 305 (1958).
453. Drury, D. R., *Am. J. Physiol.*, *131*, 536 (1940–1941).
454. Boxer, G. E., and Stetten, D., Jr., *J. Biol. Chem.*, *156*, 271 (1944).
455. Chernick, S. S., Chaikoff, I. L., Masoro, E. J. and Isaeff, E., *J. Biol. Chem.*, *186*, 527 (1950).
456. Chaikoff, I. L., *Harvey Lecture Ser.*, *47*, 99 (1953).
457. Wieland, O., Neufeldt, I., Numa, S., and Lynen, F., *Biochem. Z.*, *336*, 455 (1963).
458. Garland, P. B., and Tubbs, P. K., *Biochem. J.*, *89*, 25P (1963).
459. Tubbs, P. K., and Garland, P. B., *Biochem. J.*, *89*, 25P (1963).
460. Tubbs, P. K., and Garland, P. B., *Biochem. J.*, *93*, 550 (1964).
461. Bortz, W. M., and Lynen, F., *Biochem. Z.*, *339*, 77 (1963).
462. Garland, P. B., Shepherd, D., and Yates, D. W., *Biochem. J.*, *97*, 587 (1965).
463. Bortz, W. M., *Biochim. Biophys. Acta*, *137*, 533 (1967).

464. Maragoudakis, M. E., *Biochemistry*, *9*, 413 (1970).
465. Maragoudakis, M. E., *J. Biol. Chem.*, *244*, 5005 (1969).
466. Korchak, H. M., and Masoro, E. J., *Biochim. Biophys. Acta*, *58*, 354 (1962).
467. Allmann, D. W., Hubbard, D. D., and Gibson, D. M., *J. Lipid Res.*, *6*, 63 (1965).
468. Gellhorn, A., and Benjamin, W., *Science*, *146*, 1166 (1964).
469. Gellhorn, A., and Benjamin, W., *Advan. Enzyme Regulation*, *4*, 19 (1965).
470. Gibson, D. M., Hicks, S. E., and Allmann, D. W., *Advan. Enzyme Regulation*, *4*, 239 (1965).
471. Hicks, S. E., Allmann, D. W., and Gibson, D. M., *Biochim. Biophys. Acta*, *106*, 441 (1965).
472. Kipnis, D. M., and Kalkoff, R. J., *J. Clin. Invest.*, *44*, 1064 (1965).
473. Majerus, P. W., and Kilburn, E., *J. Biol. Chem.*, *244*, 6254 (1969).
474. Numa, S., Ringelmann, E., and Lynen, F., *Biochem. Z.*, *343*, 243 (1965).
475. Numa, S., and Ringelmann, E., *Biochem. Z.*, *343*, 258 (1965).
476. Greenspan, M., and Lowenstein, J. M., *Arch. Biochem. Biophys.*, *118*, 260 (1967).
477. Bortz, W. M., and Lynen, F., *Biochem. Z.*, *337*, 505 (1963).
478. Yugari, Y., and Matsuda, T., *J. Biochem.*, *61*, 541 (1967).
479. Yugari, Y., Matsuda, T., Miss, T., and Suda, M., *Proc. Symp. Chem. Physiol. Pathol.* (in Japanese), *3*, 43 (1963). Cited by ref. 478.
480. Levy, H. R., *Biochem. Biophys. Res. Commun.*, *13*, 267 (1963).
481. Jencks, W. P., in *Current Aspects of Biochemical Energetics*, N. O. Kaplan and E. P. Kennedy, Eds., Academic Press, New York, 1966, p. 273.
482. Greenspan, M. D., and Lowenstein, J. M., *J. Biol. Chem.*, *243*, 6273 (1968).
483. Fritz, I. B., and Hsu, M. P., *J. Biol. Chem.*, *242*, 865 (1967).
484. Fritz, I. B., and Hsu, M. P., *Biochem. Biophys. Res. Commun.*, *22*, 737 (1966).
485. Marquis, N. R., Francesconi, R. P., and Villee, C. A., *Advan. Enzyme Regulation*, *6*, 31 (1968).
486. Fang, M., and Lowenstein, J. M., *Biochem. J.*, *105*, 803 (1967).
487. Scorpio, R. M., and Masoro, E. J., *Biochem. Biophys. Res. Commun.*, *31*, 950 (1968).
488. Swanson, R. F., Curry, W. M., and Anker, H. S., *Biochim. Biophys. Acta*, *159*, 390 (1968).
489. Numa, S., in *Methods in Enzymology*, Vol. XIV, J. M. Lowenstein, Ed., Academic Press, New York, 1969, p. 9.
490. Swanson, R. F., Curry, W. M., and Anker, H. S., *Proc. Nat. Acad. Sci. U.S.*, *58*, 1243 (1967).
491. Iritani, N., Nakanishi, S., and Numa, S., *Life Sciences*, *8*, 1157 (1969).
492. Monod, J., Wyman, J., and Changeux, J.-P., *J. Mol. Biol.*, *12*, 88 (1965).
493. Knoop, F., *Beitr. Chem. Physiol. u. Path.*, *6*, 150 (1904).
494. Graffin, A. L., and Green, D. E., *J. Biol. Chem.*, *176*, 95 (1948).
495. Atchley, W. A., *J. Biol. Chem.*, *176*, 123 (1948).
496. Coon, M. J., Abrahamsen, N. S. B., and Greene, G. S., *J. Biol. Chem.*, *199*, 75 (1952).

497. Coon, M. J., and Robinson, W. G., in Methods in Enzymology, Vol. V, S. P. Colowick and N. O. Kaplan, Eds., Academic Press, New York, 1962, p. 451.
498. Coon, M. J., Robinson, W. G., and Bachhawat, B. K., in Amino Acid Metabolism, W. D. McElroy and H. B. Glass, Eds., Johns Hopkins Press, Baltimore, 1955, p. 431.
499. Kinnory, D. S., Takeda, Y., and Greenberg, D. M., J. Biol. Chem., 212, 385 (1955).
500. Phillipson, A. T., J. Exp. Biol., 23, 346 (1947).
501. Elsden, S. R., and Phillipson, A. T., Ann. Rev. Biochem., 17, 705 (1948).
502. Armstrong, D. G., and Blaxter, K. L., Brit. J. Nutr., 11, 413 (1957).
503. Annison, E. F., Hill, K. J., and Lewis, D., Biochem. J., 66, 592 (1957).
504. Reid, R. L., Hogan, J. P., and Briggs, P. K., Austral. J. Agr. Res., 8, 691 (1957).
505. Reid, R. L., Nature, 165, 448 (1950).
506. Jarrett, I. G., and Potter, B. J., Nature, 166, 515 (1950).
507. Ringer, A. I., J. Biol. Chem., 12, 511 (1912).
508. Deuel, H. J., Jr., Butts, J. S., Hallman, L. F., and Cutler, C. H., J. Biol. Chem., 112, 15 (1935-1936).
509. Buchanan, J. M., Hastings, A. B., and Nesbett, F. B., J. Biol. Chem., 150, 413 (1943).
510. Black, A. L., and Kleiber, M., J. Biol. Chem., 232, 203 (1958).
511. Hahn, A., and Haarman, W., Z. Biol., 90, 231 (1930).
512. Huennekens, F. M., Mahler, H. R., and Nordmann, J., Arch. Biochem., 30, 66 (1951).
513. Mahler, H. R., and Huennekens, F. M., Biochim. Biophys. Acta, 11, 575 (1953).
514. Lorber, V., Lifson, N., Sakami, W., and Wood, H. G., J. Biol. Chem., 183, 531 (1950).
515. Shreeve, W. W., J. Biol. Chem., 195, 1 (1952).
516. Daus, L., Meinke, M., and Calvin, M., J. Biol. Chem., 196, 77 (1952).
517. Lardy, H. A., Proc. Nat. Acad. Sci. U.S., 38, 1003 (1952).
518. Lardy, H. A., and Fischer, J., Abstr. 123rd Meeting, Amer. Chem. Soc., 10C (1953).
519. Flavin, M., Federation Proc., 14, 211 (1955).
520. Beck, W. S., and Ochoa, S., J. Biol. Chem., 232, 931 (1958).
521. Mazumder, R., Sasakawa, T., Kaziro, Y., and Ochoa, S., J. Biol. Chem., 236, PC 53 (1961).
522. Mazumder, R., and Sasakawa, T., Federation Proc., 20, 272 (1961).
523. Mazumder, R., Sasakawa, T., Kaziro, Y., and Ochoa, S., J. Biol. Chem., 237, 3065 (1962).
524. Overath, P., Kellerman, G. M., Lynen, F., Fritz, H. P., and Keller, J., Biochem. Z., 335, 500 (1962).
525. Eggerer, H., Stadtman, E. R., Overath, P., and Lynen, F., Biochem. Z., 333, 1 (1960).
526. Eggerer, H., Overath, P., Lynen, F., and Stadtman, E. R., J. Amer. Chem. Soc., 82, 2643 (1960).

527. Hegre, C. S., Miller, S. J., and Lane, M. D., Biochim. Biophys. Acta, 56, 538 (1962).
528. Kallenmeyer, R. W., and Wood, H. G., Biochemistry, 1, 1124 (1962).
529. Swick, R. W., Proc. Nat. Acad. Sci. U.S., 48, 288 (1962).
530. Phares, E. F., Long, M. V., and Carson, S. F., Biochem. Biophys. Res. Commun., 8, 142 (1962).
531. Lengyel, P., Mazumder, R., and Ochoa, S., Proc. Nat. Acad. Sci. U.S., 46, 1312 (1960).
532. Stadtman, E. R., Overath, P., Eggerer, H., and Lynen, F., Biochem. Biophys. Res. Commun., 2, 1 (1960).
533. Stjernholm, R., and Wood, H. G., Proc. Nat. Acad. Sci. U.S., 47, 303 (1961).
534. Stjernholm, R., Noble, R. E., and Kochweser, D., Biochim. Biophys. Acta, 64, 174 (1962).
535. Overath, P., Stadtman, E. R., Kellerman, G. M., and Lynen, F., Biochem. Z., 336, 77 (1962).
536. Mazumder, R., Sasakawa, T., and Ochoa, S., J. Biol. Chem., 238, 50 (1963).
537. Beck, W. S., in Methods in Enzymology, Vol. V, S. P. Colowick and N. O. Kaplan, Eds., Academic Press, New York, 1962, p. 581.
538. White, A. M., Biochem. J., 84, 41P (1962).
539. Cox, E. V., and White, A. M., Lancet, 2, 853 (1962).
540. Barness, L. A., Young, D., Mellman, W. J., Kahn, S. B., and Williams, W. J., New Eng. J. Med., 268, 144 (1963).
541. Oberholzer, V. G., Levin, B., Burgess, E. A., and Young, W. F., Arch. Disease Childhood, 42, 492 (1967).
542. Rosenberg, L. E., Lilljeqvist, A. C., and Hsia, Y. E., Science, 162, 805 (1968).
543. Mudd, S. H., Levy, H. L. and Abeles, R. H., Biochem. Biophys. Res. Commun., 35, 121 (1969).
544. Rosenberg, L. E., Lilljeqvist, A.-C., Hsia, Y. E., and Rosenbloom, F. M., Biochem. Biophys. Res. Commun., 37, 607 (1969).
545. Morrow, G., III, Barness, L. A., Cardinale, G. J., Abeles, R. H., and Flaks, J. G., Proc. Nat. Acad. Sci. U.S., 63, 191 (1969).
546. Cardinale, G. J., Carty, T. J., and Abeles, R. H., J. Biol. Chem., 245, 3771 (1970).
547. Kaziro, Y., Ochoa, S., Warner, R. C., and Chen, J.-Y., J. Biol. Chem., 236, 1917 (1961).
548. Hegre, C. S., and Lane, M. D., Biochim. Biophys. Acta, 128, 172 (1966).
549. Neujahr, H. Y., and Mistry, S. P., Acta Chem. Scand., 17, 1140 (1963).
550. Halenz, D. R., and Lane, M. D., J. Biol. Chem., 235, 878 (1960).
551. Tietz, A., and Ochoa, S., J. Biol. Chem., 234, 1394 (1959).
552. Halenz, D. R., and Lane, M. D., unpublished observations.
553. Hegre, C. S., Halenz, D. R., and Lane, M. D., J. Amer. Chem. Soc., 81, 6526 (1959).
554. Stern, J. R., Friedman, D. L., and Menon, G. K. K., Biochim. Biophys. Acta, 36, 299 (1959).
555. Stern, J. R., Friedman, D. L., and Menon, G. K. K., Arch. Biochem. Biophys., 92, 280 (1961).
556. Edwards, J. B., and Keech, D. B., Biochim. Biophys. Acta, 146, 576 (1967).

557. Edwards, J. B., doctoral thesis, submitted to the University of Adelaide (1967).
558. Greenspan, M. D., Alberts, A. W., and Vagelos, P. R., J. Biol. Chem., 244, 6477 (1969).
559. Joshi, V. C., Plate, C. A., and Wakil, S. J., J. Biol. Chem., 245, 2857 (1970).
560. Plate, C. A., Joshi, V. C., and Wakil, S. J., J. Biol. Chem., 245, 2868 (1970).
561. Phillips, G. T., Nixon, J. E., Abramovitz, A. S., and Porter, J. W., Arch. Biochem. Biophys., 138, 357 (1970).
562. Nixon, J. E., Phillips, G. T., Abramovitz, A. S., and Porter, J. W., Arch. Biochem. Biophys., 138, 372 (1970).
563. Phillips, G. T., Nixon, J. E., Dorsey, J. A., Butterworth, P. H. W., Chesterton, C. J., and Porter, J. W., Arch. Biochem. Biophys., 138, 380 (1970).
564. Neujahr, H. Y., Acta Chem. Scand., 17, 1777 (1963).
565. Edwards, J. B., and Keech, D. B., Biochem. Biophys. Acta, 159, 167 (1968).
566. Giorgio, J., and Plaut, G. W. E., Federation Proc., 26, 562 (1967).
567. Young, M. R., Tolbert, B., Valentine, R. C., Wallace, J. C., and Utter, M. F., Federation Proc., 27, 522 (1968).
568. Scrutton, M. C., Young, M. R., and Utter, M. F., J. Biol. Chem., 245, 6220 (1970).
569. Young, M. R., Tolbert, B., Valentine, R. C., and Utter, M. F., Biochemistry (1971), in press.
570. Apitz-Castro, R., Rehn, K., and Lynen, F., European J. Biochem., 16, 71 (1970).
571. Embden, G., Salomon, H., and Schmidt, F., Beitr. chem. Physiol. u. Path., 8, 129 (1906).
572. Embden, G., Beitr. Chem. Physiol. Path., 11, 348 (1908).
573. Dakin, H. D., J. Biol. Chem., 14, 321 (1913).
574. Ringer, A. I., Frankel, E. M., and Jonas, L., J. Biol. Chem., 14, 525 (1913).
575. Ikebe, Y., J. Biochem., 17, 275 (1933).
576. Edson, N. L., Biochem. J., 29, 2498 (1935).
577. Knappe, J., Schlegel, H.-G., and Lynen, F., Biochem. Z., 335, 101 (1961).
578. Mahler, H. R., and Cordes, E. H., Biological Chemistry, Harper and Row, New York, 1966, p. 691.
579. Lehninger, A. L., Biochemistry: The Molecular Basis of Cell Structure and Function, Worth, New York, 1970, p. 443.
580. McGilvery, R. W., Biochemistry: A Functional Approach, W. B. Saunders, Philadelphia, 1970, p. 376.
581. Bloch, K., J. Biol. Chem., 155, 255 (1944).
582. Coon, M. J., and Gurin, S., J. Biol. Chem., 180, 1159 (1949).
583. Zabin, I., and Bloch, K., J. Biol. Chem., 185, 117 (1950).
584. Coon, M. J., J. Biol. Chem., 187, 71 (1950).
585. Plaut, G. W. E., Proc. Soc. Exptl. Biol. Med., 78, 769 (1951).
586. Plaut, G. W. E., and Lardy, H. A., J. Biol. Chem., 192, 435 (1951).
587. Drysdale, G. R., and Lardy, H. A., J. Biol. Chem., 202, 119 (1953).
588. Knappe, J., and Lynen, F., Abstr. 4th Intern. Congr. Biochem., Vienna, 1958, 49 (1959-1960).
589. Hilz, H., Knappe, J., Ringelmann, E., and Lynen, F., Biochem. Z., 329, 476 (1958).

590. Bachhawat, B. K., Robinson, W. G., and Coon, M. J., J. Biol. Chem., 219, 539 (1956).
591. Rilling, H. C., and Coon, M. J., Federation Proc., 19, 224 (1960).
592. Rilling, H. C., and Coon, M. J., J. Biol. Chem., 235, 3087 (1960).
593. Bachhawat, B. K., Robinson, W. G., and Coon, M. J., J. Amer. Chem. Soc., 76, 3098 (1954).
594. Robinson, W. G., Bachhawat, B. K., and Coon, M. J., Federation Proc., 13, 281 (1954).
595. Coon, M. J., Federation Proc., 14, 762 (1955).
596. Bachhawat, B. K., Robinson, W. G., and Coon, M. J., J. Biol. Chem., 216, 727 (1955).
597. Danois, J., and Levitz, M., in The Metabolic Basis of Inherited Disease, J. B. Stanbury, J. B. Wyngaarden, and D. S. Fredrickson, Eds., McGraw-Hill, New York, 1966, p. 353.
598. Fischer, J. E., Proc. Soc. Exptl. Biol. Med., 88, 227 (1955).
599. Plaut, G. W. E., and Lardy, H. A., J. Biol. Chem., 186, 705 (1950).
600. Woessner, J. F., Jr., Bachhawat, B. K., and Coon, M. J., J. Biol. Chem., 233, 520 (1958).
601. Seubert, W., and Remberger, U., Biochem. Z., 338, 245 (1963).
602. Werkman, C. H., and Wood, H. G., Adv. Enzymol., 2, 135 (1942).
603. Wood, H. G., Physiol. Rev., 26, 198 (1946).
604. Wood, H. G., and Utter, M. F., Essays Biochem., 1, 1 (1965).
605. Wood, H. G., and Werkman, C. H., Biochem. J., 30, 48 (1936).
606. Wood, H. G., Stone, R. W., and Werkman, C. H., Biochem. J., 31, 349 (1937).
607. Wood, H. G., and Werkman, C. H., Biochem. J., 32, 1262 (1938).
608. Wood, H. G., and Werkman, C. H., Biochem. J., 34, 7 (1940).
609. Wood, H. G., and Werkman, C. H., Biochem. J., 34, 129 (1940).
610. Krampitz, L. O., Wood, H. G., and Werkman, C. H., J. Biol. Chem., 147, 243 (1943).
611. Elsden, S. R., Biochem. J., 32, 187 (1938).
612. Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O., J. Biol. Chem., 139, 377 (1941).
613. Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O., J. Biol. Chem., 139, 365 (1941).
614. Kalnitsky, G., and Werkman, C. H., Arch. Biochem., 4, 25 (1944).
615. Solomon, A. K., Vennesland, B., Klemperer, F. W., Buchanan, J. M., and Hastings, A. B., J. Biol. Chem., 140, 171 (1941).
616. Conant, J. B., Cramer, R. D., Hastings, A. B., Klemperer, F. W., Solomon, A. K., and Vennesland, B., J. Biol. Chem., 137, 557 (1941).
617. Vennesland, B., Solomon, A. K., Buchanan, J. M., Cramer, R. D., and Hastings, A. B., J. Biol. Chem., 142, 371 (1942).
618. Vennesland, B., Solomon, A. K., Buchanan, J. M., and Hastings, A. B., J. Biol. Chem., 142, 379 (1942).
619. Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O., J. Biol. Chem., 142, 31 (1942).
620. Evans, E. A., Jr., Slotin, L., and Vennesland, B., J. Biol. Chem., 143, 565 (1942).

621. Evans, E. A., Jr., Vennesland, B., and Slotin, L., J. Biol. Chem., 147, 771 (1943).
622. Wood, H. G., Lifson, N., and Lorber, V., J. Biol. Chem., 159, 475 (1945).
623. Anfinsen, C. B., Beloff, A., Hastings, A. B., and Solomon, A. K., J. Biol. Chem., 168, 771 (1947).
624. Crane, R. K., and Ball, E. G., J. Biol. Chem., 188, 819 (1951).
625. Weil-Malherbe, H., Biochem. J., 31, 299 (1937).
626. Krebs, H. A., and Eggleston, L. V., Biochem. J., 34, 1383 (1940).
627. Cleland, W. W., and Johnson, M. J., J. Biol. Chem., 208, 679 (1954).
628. Lewis, K. F., and Weinhouse, S., J. Amer. Chem. Soc., 73, 2906 (1951).
629. Woronick, C. L., and Johnson, M. J., J. Biol. Chem., 235, 9 (1960).
630. Bartley, W., Biochem. J., 56, 387 (1954).
631. Bandurski, R. S., and Lipmann, F., J. Biol. Chem., 219, 741 (1956).
632. Utter, M. F., and Kurahashi, K., J. Biol. Chem., 207, 821 (1954).
633. Utter, M. F., Kurahashi, K., and Rose, I. A., J. Biol. Chem., 207, 803 (1954).
634. Stanbury, S. W., and Mudge, G. H., J. Biol. Chem., 210, 949 (1954).
635. Mudge, G. H., Neuberg, H. W., and Stanbury, S. W., J. Biol. Chem., 210, 965 (1954).
636. Utter, M. F., and Kurahashi, K., J. Biol. Chem., 207, 787 (1954).
637. Graves, J. L., Vennesland, B., Utter, M. F., and Pennington, R. J., J. Biol. Chem., 233, 551 (1956).
638. Schellenberg, K. A., and Weinbach, E. C., Biochim. Biophys. Acta, 45, 593 (1960).
639. Nordlie, R. C. and Lardy, H. A., Biochem. Z., 338, 356 (1963).
640. Utter, M. F., Iowa State J. Science, 38, 97 (1963).
641. Utter, M. F., and Keech, D. B., Federation Proc., 19, 36 (1960).
642. Keech, D. B., and Utter, M. F., J. Biol. Chem., 238, 2609 (1963).
643. Utter, M. F., Keech, D. B., and Scrutton, M. C., Adv. Enzyme Regulation, 2, 49 (1964).
644. Utter, M. F., and Keech, D. B., J. Biol. Chem., 238, 2603 (1963).
645. Henning, H. V., and Seubert, W., Biochem. Z., 340, 160 (1964).
646. Ling, A.-M., and Keech, D. B., Enzymologia, 30, 367 (1966).
647. Benedict, C. R., 6th Intern. Congr. Biochem., VI, 503 (1964).
648. Losada, M., Cánovas, J. L., and Ruiz-Amil, M., Biochem. Z., 340, 60 (1964).
649. Ruiz-Amil, M., de Torrontegui, G., Palacián, E., Catalina, L., and Losada, M., J. Biol. Chem., 240, 3485 (1965).
650. Overman, S. A., and Romano, A. H., Biochem. Biophys. Res. Commun., 37, 457 (1969).
651. Fuller, R. C., and Kornberg, H. L., Biochem. J., 79, 8P (1961).
652. Hartman, R. E., J. Bact., 102, 341 (1970).
653. Bridgeland, E. S., and Jones, K. M., Biochem. J., 104, 9P (1967).
654. Cazzulo, J. J., and Sunderam, T., cited by ref. 656.
655. Bloom, S. J., and Johnson, M. J., J. Biol. Chem., 237, 2718 (1962).
656. Utter, M. F., and Scrutton, M. C., Current Topics in Cellular Regulation, Vol. 1, Academic Press, New York, 1969, p. 253.
657. Utter, M. F., Scrutton, M. C., Young, M. R., Tolbert, B., Wallace, J. C., Irias, J. J., and Valentine, R. C., Abstr. 7th Intern. Congr. Biochem., Tokyo, Symp. V—1, 4, II, 247 (1967).

658. Scrutton, M. C., Olmsted, M. R., and Utter, M. F., in Methods in Enzymology, Vol. 13, J. M. Lowenstein, Ed., Academic Press, New York, 1969, p. 235.
659. Grüninger, P., and Scrutton, M. C., Federation Proc., 29, 765 (1970).
660. Seubert, W., and Weicker, H., in Methods in Enzymology, Vol. 13, J. M. Lowenstein, Ed., Academic Press, New York, 1969, p. 258.
661. Cazzulo, J. J., and Stoppani, A. O. M., Arch. Biochem. Biophys., 121, 596 (1967).
662. Scrutton, M. C., and Mildvan, A. S., in CO_2: Chemical, Biochemical, and Physiological Aspects, R. E. Forster, J. T. Edsall, A. B. Otis and F. J. W. Roughton, Eds., NASA SP-188, U.S. Government Printing Office, Washington, D.C., 1969, p. 207.
663. Scrutton, M. C., and Mildvan, A. S., Biochemistry, 7, 1490 (1968).
664. Scrutton, M. C., and Mildvan, A. S., Arch. Biochem. Biophys., 140, 131 (1970).
665. Scrutton, M. C., Federation Proc., 28, 534 (1969).
666. Mildvan, A. S., and Scrutton, M. C., Abstr. 7th Intern. Congr. Biochem., Tokyo, 787 (1967).
667. Scrutton, M. C., and Utter, M. F., J. Biol. Chem., 240, 1 (1965).
668. Klotz, I. M., Langerman, N. R., and Darnall, D. W., Ann. Rev. Biochem., 39, 25 (1970).
669. Irias, J. J., Olmsted, M. R., and Utter, M. F., Biochemistry, 8, 5136 (1969).
670. Irias, J. J., Federation Proc., 24, 411 (1965).
671. Scrutton, M. C., and Utter, M. F., J. Biol. Chem., 242, 1723 (1967).
672. Fung, C. H., and Utter, M. F., Federation Proc., 29, 542 (1970).
673. Cooper, T. G., and Benedict, C. R., Biochemistry, 7, 3032 (1968).
674. Cooper, T. G., and Benedict, C. R., Biochem. Biophys. Res. Commun., 22, 285 (1966).
675. Cazzulo, J. J., and Stoppani, A. O. M., Biochim. Biophys. Acta, 100, 276 (1965).
676. Ashworth, J. M., and Kornberg, H. L., Proc. Roy. Soc., B165, 179 (1966).
677. Kornberg, H. L., in Essays in Biochemistry, Vol. II, P. N. Campbell and G. D. Greville, Eds., Academic Press, New York, 1966, p. 1.
678. Palacián, E., de Torrontegui, G., and Losada, M., Biochem. Biophys. Res. Commun., 24, 644 (1966).
679. Cazzulo, J. J., and Stoppani, A. O. M., Arch. Biochem. Biophys., 127, 563 (1968).
680. Holden, J. T., in Amino Acid Pools, J. T. Holden, Ed., Elsevier, Amsterdam, 1962, p. 73. Cited by ref. 679.
681. Scrutton, M. C., and Utter, M. F., Ann. Rev. Biochem., 37, 249 (1968).
682. Krebs, H. A., Advan. Enzyme Regulation, 1, 385 (1963).
683. Lardy, H. A., Harvey Lectures, 60, 261 (1966).
684. Krebs, H. A., Proc. Roy. Soc. (London), Ser. B, 159, 545 (1964).
685. Utter, M. F., Ann. N.Y. Acad. Sci., U.S., 72, 451 (1959).
686. Böttger, I., Wieland, O., Brdiczka, D., and Pette, D., European J. Biochem., 8, 113 (1969).
687. Lardy, H. A., Paetkau, V., and Walter, P., Proc. Nat. Acad. Sci. U.S., 53, 1410 (1965).
688. Walter, P., Paetkau, V., and Lardy, H. A., J. Biol. Chem., 241, 2523 (1966).

689. Williamson, J. R., Kreisberg, R. A., and Felts, P. W., Proc. Nat. Acad. Sci. U.S., 56, 247 (1966).
690. Haynes, R. C., Jr., J. Biol. Chem., 240, 4103 (1965).
691. Nordie, R. C., and Lardy, H. A., Federation Proc., 21, 229 (1962).
692. Krebs, H. A., Bennett, D. A. H., DeGasquet, P., Gascoyne, T., and Yoshida, T., Biochem. J., 86, 22 (1963).
693. Freedman, A. C., and Kohn, L., Science, 145, 58 (1964).
694. Exton, J. H., Jefferson, L. S., Butcher, R. W., and Park, C. R., Amer. J. Medicine, 40, 709 (1966).
695. Wagle, S. R., Biochem. Biophys. Acta, 97, 142 (1965).
696. Wagle, S. R., Biochem. Biophys. Res. Commun., 14, 533 (1964).
697. Wagle, S. R., and Ashmore, J., Biochim. Biophys. Acta, 74, 564 (1963).
698. Wagle, S. R., and Ashmore, J., J. Biol. Chem., 238, 17 (1963).
699. Wagle, S. R., Diabetes, 15, 19 (1966).
700. Prinz, W., and Seubert, W., Biochem. Biophys. Res. Commun., 16, 582 (1964).
701. Koepf, G. F., Horn, H. W., Gemmill, C. L., and Thorn, G. W., Amer. J. Physiol., 135, 175 (1941).
702. Welt, I. D., Stetten, D., Jr., Ingle, D. J., and Morley, E. H., J. Biol. Chem., 197, 57 (1952).
703. Henning, H. V., Seifort, I., and Seubert, W., Biochim. Biophys. Acta, 77, 345 (1963).
704. Seubert, W., Henning, H. V., Schoner, W., and Läge, M., Advan. Enzyme Regulation, 6, 153 (1968).
705. Henning, H. V., Huth, W., and Seubert, W., Biochem. Biophys. Res. Commun., 17, 496 (1964).
706. Haynes, R. C., Jr., Advan. Enzyme Regulation, 3, 111 (1965).
707. Eisenstein, A. B., Advan. Enzyme Regulation, 3, 121 (1965).
708. Weber, G., Singhal, R. L., and Srivastava, S. K., Advan. Enzyme Regulation, 3, 43 (1965).
709. Ray, P. D., Foster, D. O., and Lardy, H. A., J. Biol. Chem., 239, 3396 (1964).
710. Haynes, R. C., Jr., Endocrinology, 71, 399 (1962).
711. Eisenstein, A. B., Berg, E., Goldenberg, D., and Jensen, B., Endocrinology, 74, 123 (1964).
712. Exton, J. H., and Park, C. R., Federation Proc., 24, 537 (1965).
713. Uete, T., and Ashmore, J., J. Biol. Chem., 238, 2806 (1963).
714. Henning, H. V., Stumpf, B., Ohly, B., and Seubert, W., Biochem. Z., 344, 274 (1966).
715. Landau, B. R., Mahler, R., Ashmore, J., Elwyn, D., Hastings, A. B., and Zottu, S., Endocrinology, 70, 47 (1962).
716. Williamson, J. R., Advan. Enzyme Regulation, 5, 229 (1967).
717. Struck, E., Ashmore, J., and Wieland, O., Advan. Enzyme Regulation, 4, 219 (1966).
718. Williamson, J. R., Biochem. J., 101, 11c (1966).
719. Garcia, A., Williamson, J. R., and Cahill, G. F., Jr., Diabetes, 15, 188 (1966).
720. Bondy, P. K., Bloom, W. L., Whitner, V. S., and Farrar, B., J. Clin. Invest., 28, 1126 (1949).

721. Berry, M. N., Biochem. J., 95, 587 (1965).
722. Barritt, G. J., Keech, D. B., and Ling, A.-M., Biochem. Biophys. Res. Commun., 24, 476 (1966).
723. Krebs, H. A., Speake, R. N., and Hems, R., Biochem. J., 94, 712 (1965).
724. Wieland, O., and Weiss, L., Biochem. Biophys. Res. Commun., 13, 26 (1963).
725. Wieland, O., Weiss, L., and Eger-Neufeldt, I., Advan. Enzyme Regulation, 2, 85 (1964).
726. Krebs, H. A., Newsholme, E. A., Speake, R., Gascoyne, T., and Lund, P., Advan. Enzyme Regulation, 2, 71 (1964).
727. Williamson, J. R., Peterson, M., and Rostand, S., Federation Proc., 29, 735 (1970).
728. Walker, J. B., Proc. Nat. Acad. Sci. U.S., 38, 561 (1952).
729. Hattori, A., J. Biochem., 44, 253 (1957).
730. Kating, H., Biochem. Z., 335, 366 (1961-2).
731. Domnas, A., J. Biochem., 52, 149 (1962).
732. Ellner, P. D., and Steers, E., Arch. Biochem. Biophys., 59, 534 (1955).
733. Roon, R. J., and Levenberg, B., Federation Proc., 28, 668 (1969).
734. Schultz, A. S., Atkin, L., and Frey, C. N., J. Biol. Chem., 135, 267 (1940).
735. DiCarlo, F. J., Schultz, A. S., and Kent, A. M., J. Biol. Chem., 199, 333 (1952).
736. DiCarlo, F. J., Schultz, A. S., and Kent, A. M., Arch. Biochem. Biophys., 44, 468 (1953).
737. Rognes, S. E., Roon, R. J., and Levenberg, B., FEBS Lett., 3, 233 (1969).
738. Roon, R. J., and Levenberg, B., J. Biol. Chem., 243, 5213 (1968).
739. Roon, R. J., and Levenberg, B., in Methods in Enzymology, Vol. 17, C. W. Tabor and H. Tabor, Eds., Academic Press, New York, 1970, in press.
740. Whitney, P. A., and Cooper, T. G., Biochem. Biophys. Res. Commun., 40, 814 (1970).
741. Allen, S. H. G., Kellermeyer, R. W., Stjernholm, R. L., and Wood, H. G., J. Bact., 87, 171 (1964).
742. Wood, H. G., and Werkman, C. H., J. Bact. 30, 332 (1935).
743. Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O., Proc. Soc. Exptl. Biol. Med., 46, 313 (1941).
744. Wood, H. G., Kulka, R. G., and Edson, N. L., Biochem. J., 63, 177 (1956).
745. Anthony, D. S., Carson, S. F., Kuna, M., and Phares, E. F., Federation Proc., 8, 5 (1949).
746. Barban, S., and Ajl, S. J., J. Biol. Chem., 192, 63 (1951).
747. Leaver, F. W., Wood, H. G., and Stjernholm, R., J. Bact., 70, 521 (1955).
748. Wood, H. G., Stjernholm, R., and Leaver, F. W., J. Bact., 72, 142 (1956).
749. Pomerantz, S. H., and Wood, H. G., J. Biol. Chem., 231, 519 (1958).
750. Pomerantz, S. H., J. Biol. Chem., 231, 505 (1958).
751. Carson, S. F., and Ruben, S., Proc. Nat. Acad. Sci. U.S., 26, 422 (1940).
752. Carson, S. F., Foster, J. W., Ruben, S., and Barker, H. A., Proc. Nat. Acad. Sci. U.S., 27, 229 (1941).
753. Delwiche, E. A., J. Bact., 56, 811 (1948).
754. Delwiche, E. A., Phares, E. F., and Carson, S. F., J. Bact., 71, 598 (1956).
755. Phares, E. F., Delwiche, E. A., and Carson, S. F., J. Bact., 71, 604 (1956).

756. Delwiche, E. A., Phares, E. F., and Carson, S. F., *Federation Proc.*, *12*, 194 (1953).
757. Wood, H. G., and Leaver, F. W., *Biochim. Biophys. Acta*, *12*, 207 (1953).
758. Delwiche, E. A., Phares, E. F., and Carson, S. F., *Federation Proc.*, *13*, 198 (1954).
759. Sprecher, M., Clark, M. J., and Sprinson, D. B., *J. Biol. Chem.*, *241*, 872 (1966).
760. Allen, S. H. G., Kellermeyer, R., Stjernholm, R., Jacobson, B., and Wood, H. G., *J. Biol. Chem.*, *238*, 1637 (1963).
761. Kellermeyer, R. W., Allen, S. H. G., Stjernholm, R., and Wood, H. G., *J. Biol. Chem.*, *239*, 2562 (1964).
762. Wood, H. G., Jacobson, B., Gerwin, B. I., and Northrop, D. B., in *Methods in Enzymology*, Vol. 13, J. M. Lowenstein, Ed., Academic Press, New York, 1969, p. 215.
763. Wood, H. G., Allen, S. H. G., Stjernholm, R., and Jacobson, B., *J. Biol. Chem.*, *238*, 547 (1963).
764. Ahmad, F., Jacobson, B., Lygre, D., and Wood, H. G., *Federation Proc.*, *29*, 333 (1970).
765. Gerwin, B. I., Jacobson, B., Lygre, D. G., and Wood, H. G., *Federation Proc.*, *28*, 853 (1969).
766. Foubert, E. L., Jr., and Douglas, H. C., *J. Bact.*, *56*, 25 (1948).
767. Foubert, E. L., Jr., and Douglas, H. C., *J. Bact.*, *56*, 35 (1948).
768. Whiteley, H. R., *J. Amer. Chem. Soc.*, *75*, 1518 (1953).
769. Whiteley, H. R., *Proc. Nat. Acad. Sci. U.S.*, *39*, 772 (1953).
770. Whiteley, H. R., *Proc. Nat. Acad. Sci. U.S.*, *39*, 779 (1953).
771. Galivan, J. H., and Allen, S. H. G., *Federation Proc.*, *26*, 562 (1967).
772. Phares, E. F., and Carson, S. F., *Bacteriol. Proc.*, *154* (1960). Cited by 116.
773. Allen, S. H. G., and Galivan, J. H., Jr., *Federation Proc.*, *24*, 531 (1965).
774. Parsons, H. T., and Kelly, E., *Am. J. Physiol.*, *104*, 150 (1933).
775. György, P., Rose, C. S., and Tomarelli, R., *J. Biol. Chem.*, *144*, 169 (1942).
776. Wei, R.-D., and Wright, L. D., *Proc. Soc. Exptl. Biol. Med.*, *117*, 341 (1964).
777. Boas, M. A., *Biochem. J.*, *18*, 422 (1924).
778. Parsons, H. T., *J. Biol. Chem.*, *90*, 351 (1931).
779. Eakin, R. E., McKinley, W. A., and Williams, R. J., *Science*, *92*, 224 (1940).
780. György, P., and Rose, C. S., *Proc. Soc. Exptl. Biol. Med.*, *53*, 55 (1943).
781. György, P., and Rose, C. S., *Science*, *94*, 261 (1941).
782. Eakin, R. E., Snell, E. E., and Williams, R. J., *J. Biol. Chem.*, *136*, 801 (1940).
783. György, P., Rose, C. S., Eakin, R. E., Snell, E. E., and Williams, R. J., *Science*, *93*, 477 (1941).
784. Hertz, R., and Sebrell, W. H., *Science*, *96*, 257 (1942).
785. Chaiet, L., and Wolf, F. J., *Arch. Biochem. Biophys.*, *106*, 1 (1964).
786. Tausig, F., and Wolf, F. J., *Biochem. Biophys. Res. Commun.*, *14*, 205 (1964).
787. Fraps, R. M., Hertz, R., and Sebrell, W. H., *Proc. Soc. Exptl. Biol. Med.*, *52*, 140 (1943).

788. Hertz, R., Fraps, R. M., and Sebrell, W. H., *Science*, *100*, 35 (1944).
789. Hertz, R., Dhyse, F. G., and Tullner, W. W., *Endocrinology*, *45*, 461 (1949).
790. Hertz, R., Fraps, R. M., and Sebrell, W. H., *Proc. Soc. Exptl. Biol. Med.*, *52*, 142 (1943).
791. Woolley, D. W., and Longsworth, L. G., *J. Biol. Chem.*, *142*, 285 (1942).
792. Pennington, D., Snell, E. E., and Eakin, R. E., *J. Amer. Chem. Soc.*, *64*, 469 (1942).
793. Fraenkel-Conrat, H., Ward, W., Snell, N. S., and Ducay, E. D., *J. Amer. Chem. Soc.*, *72*, 3826 (1950).
794. Fraenkel-Conrat, H., Snell, N. S., and Ducay, E. D., *Arch. Biochem. Biophys.*, *39*, 97 (1952).
795. DeLange, R. J., *Federation Proc.*, *28*, 343 (1969).
796. DeLange, R. J., *J. Biol. Chem.*, *245*, 907 (1970).
796a. DeLange, R. J., and Huang, T.-S., *Federation Proc.*, *29*, 534 (1970).
796b. Huang, T.-S., and DeLange, R. J., *J. Biol. Chem.*, *246*, 686 (1971).
796c. DeLange, R. J., and Huang, T-S., *J. Biol. Chem.*, *246*, 698 (1971).
797. Green, N. M., and Toms, E. J., *Biochem. J.*, *118*, 67 (1970).
798. Green, N. M., *Biochem. J.*, *89*, 609 (1963).
799. Green, N. M., *Biochem. J.*, *92*, 16c (1964).
800. Fraenkel-Conrat, H., and Potter, R. R., *Biochim. Biophys. Acta*, *9*, 557 (1952).
801. Green, N. M., *Nature*, *217*, 254 (1968).
802. Green, N. M., *Biochem. J.*, *104*, 64P (1967).
803. Green, N. M., and Joynson, M. A., *Biochem. J.*, *118*, 71 (1970).
804. Green, N. M., *Biochem. J.*, *90*, 564 (1964).
805. Green, N. M., *Biochem. J.*, *101*, 774 (1966).
806. Green, N. M., *Biochem. J.*, *89*, 599 (1963).
807. Green, N. M., *Biochem. J.*, *94*, 23c (1965).
808. Green, N. M., *Biochem. Biophys. Acta*, *59*, 244 (1962).
809. Green, N. M., and Ross, M. E., *Biochem. J.*, *110*, 59 (1968).
810. Lichstein, H. C., and Birnbaum, J., *Biochem. Biophys. Res. Commun.*, *20*, 41 (1965).
811. Green, N. M., and Melamed, M. D., *Biochem. J.*, *100*, 614 (1966).
812. Green, N. M., *Methods in Enzymology*, in press.
813. du Vigneaud, V., Dittmer, K., Hofmann, K., and Melville, D. B., *Proc. Soc. Exptl. Biol. Med.*, *50*, 374 (1942).
814. Dittmer, K., du Vigneaud, V., György, P., and Rose, C. S., *Arch. Biochem.*, *4*, 229 (1944).
815. Wright, L. D., Skeggs, H. R., and Cresson, E. L., *Proc. Soc. Exptl. Biol. Med.*, *64*, 150 (1947).
816. Wright, L. D., and Skeggs, H. R., *Arch. Biochem.*, *12*, 27 (1947).
817. Wright, L. D., Valentik, K. A., Nepple, H. M., Cresson, E. L., and Skeggs, H. R., *Proc. Soc. Exptl. Biol. Med.*, *74*, 273 (1950).
818. Wright, L. D., Skeggs, H. R., and Cresson, E. L., *J. Amer. Chem. Soc.*, *73*, 4144 (1951).
819. Wei, R.-D., and Wright, L. D., *Proc. Soc. Exptl. Biol. Med.*, *117*, 17 (1964).

JOEL MOSS AND M. DANIEL LANE

820. Melamed, M. D., and Green, N. M., *Biochem. J.*, *89*, 591 (1963).
821. Cuatrecasas, P., and Wilchek, M., *Biochem. Biophys. Res. Commun.*, *33*, 235 (1968).
822. Thoma, R. W., and Peterson, W. H., *J. Biol. Chem.*, *210*, 569 (1954).
823. Koivusalo, M., Elorriaga, C., Kaziro, Y., and Ochoa, S., *J. Biol. Chem.*, *238*, 1038 (1963).
824. Knappe, J., Brümmer, W., and Biederbick, K., *Biochem. Z.*, *338*, 599 (1963).

SOME QUESTIONS ABOUT THE STRUCTURE AND ACTIVITY OF AMINOACYL-tRNA SYNTHETASES*

By ALAN H. MEHLER and KALPANA CHAKRABURTTY,

Milwaukee, Wisconsin

CONTENTS

Recombinant Core Streptavidins

A MINIMUM-SIZED CORE STREPTAVIDIN HAS ENHANCED STRUCTURAL STABILITY AND HIGHER ACCESSIBILITY TO BIOTINYLATED MACROMOLECULES*

Takeshi Sano, Mark W. Pandori‡, Xiaomin Chen§, Cassandra L. Smith, and Charles R. Cantor¶

From the Center for Advanced Biotechnology and Departments of Biomedical Engineering and Pharmacology, Boston University, Boston, Massachusetts 02215 and the §Department of Molecular and Cell Biology, University of California, Berkeley, California 94720

Two recombinant core streptavidins were designed and characterized to understand the role of the terminal sequences, present in naturally truncated core streptavidins, in the properties of streptavidin. One recombinant core streptavidin, Stv-25, has an amino acid sequence very similar to natural core streptavidins. The other recombinant molecule, Stv-13, has further truncation of the terminal residues and consists essentially of only the β-barrel structure characteristic of streptavidin. These recombinant core streptavidins are tetrameric and bind four biotins/molecule, as does natural streptavidin. The solubility characteristics of Stv-13, determined by varying the concentration of ammonium sulfate or ethanol, were almost the same as those of Stv-25 and natural core streptavidin. However, Stv-13 showed an enhanced structural stability compared with Stv-25 and natural core streptavidin. For example, Stv-13 retained greater than 80% of its biotin binding ability after incubation in 6 M guanidine hydrochloride at pH 1.5, under which conditions, Stv-25 and natural core streptavidin retained only about 20% of their biotin binding ability. In addition, Stv-13 showed higher accessibility to biotinylated DNA than natural core streptavidin. Apparently, the terminal regions, present on the surface of natural core streptavidin, spatially hinder biotinylated macromolecules from approaching the biotin binding sites.

Streptavidin, a protein produced by *Streptomyces avidinii*, binds D-biotin with a remarkably high affinity ($K_d \sim 10^{-15}$ M) (1–4). This extremely tight biotin binding affinity has made the streptavidin-biotin system a powerful biological tool in a variety of bioanalytical applications (5, 6). Streptavidin is generally isolated from culture media of *S. avidinii*. Such streptavidin molecules usually have truncated terminal sequences due to postsecretory cleavage of the terminal regions, which are highly susceptible to proteolysis (4, 7–10). Nontruncated or only partially truncated streptavidins tend to form higher order aggregates and thus have poor solubility. In contrast, fully truncated streptavidin, termed natural core streptavidin, is

free from aggregate formation and shows high solubility. However, the terminal sequences of natural core streptavidin often differ from preparation to preparation, and this heterogeneity can be seen even within single tetrameric molecules (9). Thus, many commercial preparations are treated with proteinases to further truncate the terminal sequences and to maximize the homogeneity of the terminal structure.

In addition to numerous biotechnological applications, streptavidin generates considerable protein chemical interest, particularly as an attractive model for studying macromolecule-ligand interactions (11–18). The determination of the three-dimensional structure of core streptavidin by x-ray crystallography (19, 20) considerably expanded the understanding of the structural characteristics of this protein at the molecular level. However, no precise information about the structure of the chain termini was obtained in these studies because of the weak densities seen for these regions in the electron density maps. This indicates that the terminal regions, located on the surface of the molecule, are rather disordered and flexible (20) and that the terminal sequences have little contribution to the fundamental properties of streptavidin, such as formation of the stable β-barrel structure and biotin binding. Apparently, these disordered structures are also responsible for the high proteinase susceptibility of the terminal regions.

The objective of the present work was to produce, by genetic engineering, core streptavidin that has a homogeneous structure. Such structurally homogeneous streptavidin molecules should be very useful in obtaining deeper understanding of the properties and structural characteristics of streptavidin. We were particularly interested in designing a minimum sized core streptavidin, which might have enhanced properties over natural core streptavidins due to the lack of nonfunctional terminal residues. In this work, two recombinant core streptavidins were designed and produced. One recombinant core streptavidin has a structure very similar to natural core streptavidins; the other has further truncation of the terminal sequences, which have no apparent function. These recombinant core streptavidins were characterized to understand the roles of the terminal regions in the properties of streptavidin.

EXPERIMENTAL PROCEDURES

Construction of Expression Vectors—Expression vectors for recombinant core streptavidins were constructed using standard techniques (21). Oligonucleotide-directed *in vitro* mutagenesis (22) was used to introduce mutations into the coding sequence for streptavidin.

Two expression vectors, pTSA-13 and pTSA-25 (Fig. 1), were constructed using a cloned natural streptavidin gene (7) as the starting material. pTSA-13 carries a DNA sequence encoding amino acid residues 16–133 of mature streptavidin (7), while pTSA-25 encodes amino acid residues 14–138. The coding sequences were cloned into pET-3a (23, 24) under the control of the Φ*10* promoter of bacteriophage T7.

Expression and Purification of Recombinant Core Streptavidins—Expression of each recombinant core streptavidin was carried out by the

‡ Present address: Depts. of Pathology and Medicine, University of California at San Diego, La Jolla, CA 92039.

¶ To whom correspondence should be addressed: Center for Advanced Biotechnology, Boston University, 36 Cummington St., Boston, MA 02215. Tel.: 617-353-8500; Fax: 617-353-8501.

Mature, full-length streptavidin (159 residues, 16.5 kDa)

```
1                                    21                              138                                        159
DPSKD SKAQV SAAEA GITGT[                                              ]TKVKP SAASI DAAKK AGVNN GNPLD AVQQ
```

Natural core streptavidin (127 residues, 13.3 kDa)

```
13      21                              130     139
AEA GITGT[                              ]TKVKP SAAS
```

Stv-25 (126 residues, 13.2 kDa)

```
14      21                              130     138
M EA GITGT[                             ]TKVKP SAA
```

Stv-13 (119 residues, 12.6 kDa)

```
16      21                              130 133
M GITGT[                                ]TKV
```

FIG. 1. **Schematic illustration of the structures of various streptavidin constructs.** The amino acid sequence is based on Argaraña *et al.* (7). Single-letter amino acid codes are used to indicate terminal sequences. A *box* represents the sequence from Thr-20 to Phe-130. Stv-13 and Stv-25 are recombinant core streptavidins designed in this work. The structure of natural core streptavidin, obtained from Boehringer Mannheim, is from Bayer *et al.* (10).
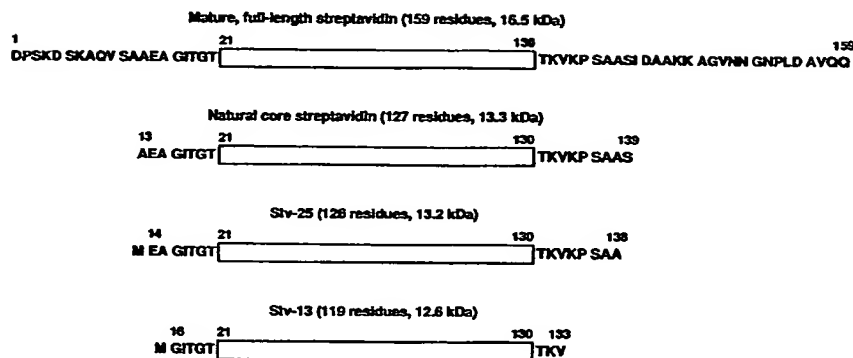
T7 expression system using BL21(DE3)(pLysE) (24) carrying an expression vector, as described previously (25–27).

Purification of Stv-13 and Stv-25 was carried out by the method described previously (25–27), including 2-iminobiotin affinity chromatography (28). BL21(DE3)(pLysE) carrying pTSA-13 or pTSA-25, which had been incubated for 4 h after induction, was used as the source.

Determination of Solubility Characteristics—The solubility characteristics of recombinant core streptavidins, Stv-13 and Stv-25, without or with biotin were determined by varying concentrations of ammonium sulfate or ethanol. Natural core streptavidin (Boehringer Mannheim) was also analyzed for comparison.

For analysis in the absence of biotin, the concentration of each core streptavidin was adjusted to 5.7 nmol of subunit/ml in Tris-buffered saline (150 mM NaCl, 20 mM Tris-Cl (pH 7.4), 0.02% NaN$_3$). This corresponds to 72 µg/ml for Stv-13, and 76 µg/ml for Stv-25 and natural core streptavidin. To 100 µl of this protein solution, 1.1 ml of an appropriate ammonium sulfate solution in Tris-buffered saline was added to adjust the final concentration of ammonium sulfate (final streptavidin concentration, 0.48 nmol of subunit/ml). The mixture was allowed to stand at 30 °C for 30 min and centrifuged at 2,200 × g for 20 min. The amount of soluble streptavidin in the supernatant fraction was determined by biotin binding assays described below. The fraction of original streptavidin remaining in the supernatant is defined as the relative solubility.

For analysis in the presence of biotin, the procedure was almost the same as above, but the biotin binding sites of each core streptavidin were saturated by adding an equimolar amount of D-[*carbonyl*-^{14}C]biotin (53 mCi/mmol; Amersham Corp.) prior to the addition of an ammonium sulfate solution. The amount of soluble streptavidin in the final supernatant was estimated from the radioactivity derived from bound biotin, determined by liquid scintillation counting.

When ethanol was used, the procedures were essentially the same as those used with ammonium sulfate, but the following modifications were made. After the addition of ethanol, the final volume was adjusted to 1.2 ml by the addition of an appropriate ethanol solution to make the final protein concentration constant for all samples. After incubation at 30 °C for 30 min, centrifugation was performed at 13,000 × g for 20 min.

Stability against Denaturation by GdnHCl[1]—The structural stability of core streptavidins was estimated from the biotin binding ability after incubation in GdnHCl solutions at pH 7.4 or pH 1.5. Each of Stv-13, Stv-25, and natural core streptavidin was incubated at 22 °C for 10 min in 500 µl of an appropriate GdnHCl solution (final GdnHCl concentration, 0–6.0 M) at a protein concentration of 270 pmol subunits/ml (1.7 µg/ml for Stv-13 and 1.8 µg/ml for Stv-25 and natural core streptavidin). Then, 1.4 µl (680 pmol) of D-[*carbonyl*-^{14}C]biotin was added to each solution. The mixture was incubated at 22 °C for 10 min, and streptavidin-biotin complexes were separated from free unbound biotin using PD-10 columns (Pharmacia Biotech Inc.), which had been equilibrated with the same GdnHCl solution. The amount of radioactive biotin remaining bound to streptavidin was determined by liquid scintillation counting.

Binding Ability for Biotinylated DNA—The binding ability of two core streptavidin species, Stv-13 and natural core streptavidin, for biotinylated DNA was determined by using a 3179-base pair linear

double-stranded DNA target in which one of the 3′ termini contains biotin. This target DNA was prepared by using an *Acc*I-*Hin*dIII fragment of the plasmid pGEM-3Zf(+) (Promega). Biotin was incorporated into the *Hin*dIII terminus by fill-in reactions in the presence of a biotinylated deoxynucleotide analog, biotin-14-dATP (Life Technologies, Inc.), as described earlier (29). Core streptavidin and the biotinylated target DNA were mixed at various ratios in 10 mM Tris-Cl (pH 8.0), 0.1 mM EDTA, and the mixtures were incubated at 37 °C for 90 min followed by electrophoretic separation on 1% agarose gels. DNA was stained with ethidium bromide.

Other Methods—Gel filtration chromatography was carried out at room temperature (22 °C) using a Sephacryl S-300 HR column (1.6 × 85 cm; Pharmacia), as described previously (26, 30). Biotin binding ability was determined by gel filtration (31) using D-[*carbonyl*-^{14}C]biotin and PD-10 columns. SDS-PAGE (32) was carried out using 15% polyacrylamide gels. Proteins were stained with Coomassie Brilliant Blue R-250. The concentration of each streptavidin preparation was determined from the absorbance at 280 nm using the following extinction coefficients ($E^{0.1\%}_{280 nm}$): Stv-13, 3.55; Stv-25, 3.35; natural core streptavidin, 3.35 (33).

RESULTS AND DISCUSSION

Design of Recombinant Core Streptavidins—Although mature streptavidin has 159 amino acids/subunit (7), such full-length, nontruncated molecules can rarely be seen under the conditions generally used for the culture of *S. avidinii*. This is due to very high susceptibility of the terminal regions of the full-length molecule to proteolysis. Such full-length and only partially truncated molecules tend to form higher-order aggregates and have poor solubility (4, 7–10), although the biological reason why *S. avidinii* produces streptavidin with poor solubility and a tendency to aggregate is unknown. For these reasons, such streptavidins are not useful in bioanalytical applications. In contrast, fully truncated core streptavidins have high solubility and show little tendency to aggregate (4, 7–10). Thus, many commercial streptavidin preparations include proteinase treatment to ensure full truncation of the terminal sequences. However, the variable performance seen for natural core streptavidin preparations may be attributable to incomplete truncation of the terminal sequences and residual proteolytic activity, resulting from proteinase treatment (34).

The objective of the present work was to design recombinant core streptavidin with a homogeneous structure. A particular motivation was derived from the fact that the structural heterogeneity of natural core streptavidins reduces the resolution obtainable in x-ray diffraction studies on streptavidin crystals (19, 20). We were particularly interested in designing a minimum sized core streptavidin, which might have enhanced properties because of the removal of any nonfunctional terminal sequences that are located on the surface of the molecules.

In this work, two recombinant core streptavidins were designed (Fig. 1). One recombinant core streptavidin, Stv-25, has amino acid residues from Glu-14 to Ala-138 plus a methionine

---

[1] The abbreviations used are: GdnHCl, guanidine hydrochloride; PAGE, polyacrylamide gel electrophoresis.
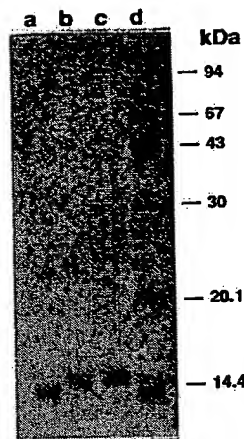
FIG. 2. **SDS-PAGE analysis of purified core streptavidins.** *Lane a,* Stv-13; *lane b,* Stv-25; *lane c,* natural core streptavidin; *lane d,* molecular mass standard proteins (Pharmacia). Approximately 2 μg of protein was applied to each lane of a 15% polyacrylamide gel. Proteins were stained with Coomassie Brilliant Blue.

residue at the N terminus, derived from a translation initiation codon, and thus has a structure very similar to natural core streptavidins. Other groups have also produced recombinant core streptavidins (34–36), which are very similar to Stv-25.

The other recombinant core streptavidin, Stv-13, has further truncation of the terminal sequences and consists of amino acid residues from Gly-16 to Val-133 plus a methionine residue at the N terminus. Previous crystallographic studies on streptavidin using natural core streptavidin were able to refine the molecular structure only from Ala-13 or Glu-14 to Val-133 (19, 20), which corresponds almost perfectly to the stable β-barrel structure consisting of the sequence from Gly-19 to Val-133. This implies that the terminal regions of natural core streptavidins have little contribution to the fundamental properties of streptavidin, which should not be altered by the further truncation of the terminal sequences made on Stv-13.

Expression and Purification of Recombinant Core Streptavidins—Expression of Stv-13 and Stv-25 was carried out using the T7 expression system, which allows efficient expression of various recombinant streptavidin constructs (25–27, 33–38). Stv-13 was expressed very efficiently in *Escherichia coli* as were other recombinant streptavidin constructs (25–27, 37). In contrast, the expression efficiency of Stv-25 was considerably lower. This is probably caused by codons for the terminal sequences present in Stv-25 (but absent in Stv-13) that occur at low frequencies in highly expressed *E. coli* genes. A similar observation was reported with another recombinant core streptavidin (35), where the expression efficiency in *E. coli* was rather low with the natural streptavidin gene but significantly improved by using a synthetic gene containing codons observed in highly expressed *E. coli* genes.

Expressed Stv-13 and Stv-25 were purified to homogeneity (Fig. 2) using a simple procedure that includes 2-iminobiotin affinity chromatography. SDS-PAGE analysis of purified proteins shows a clear difference in subunit molecular mass (650 Da) between Stv-13 and Stv-25. Natural core streptavidin obtained from Boehringer Mannheim also showed a single band on SDS-PAGE, and its migration was very similar to that of Stv-25. Although no terminal sequences were determined on the particular batch, this natural core streptavidin is likely to consist of amino acid residues 13–139, as shown by the terminal sequence analysis of the protein obtained from the same source (10), because the identity of the terminal residues is determined primarily by the proteinase treatment used.

Each of Stv-13, Stv-25, and natural core streptavidin bound greater than 0.96 molecules of biotin per subunit, indicating that these molecules have full biotin binding ability. Gel filtration chromatography using Sephacryl S-300HR showed that each of these core streptavidins is tetrameric and free from aggregate formation.

Solubility Characteristics of Core Streptavidins—Significant differences are observed in solubility characteristics of core streptavidin and full-length or only partially truncated streptavidin (4, 7–10). These differences suggest that the terminal regions primarily determine the solubility characteristics of streptavidin. To understand the effect of terminal sequences remaining in natural core streptavidin on the solubility characteristics, the solubility of each core streptavidin species with and without biotin was investigated by varying the concentration of ammonium sulfate or ethanol.

The relative solubility of the three core streptavidins, Stv-13, Stv-25, and natural core streptavidin, as the concentration of ammonium sulfate was altered, showed biphasic changes (Fig. 3, *A* and *B*); the solubility decreased sharply with increasing concentrations of ammonium sulfate up to 50% saturation and then increased with further increases in ammonium sulfate concentration. In the absence of biotin, Stv-13 showed slightly lower solubility than Stv-25 and natural core streptavidin at ammonium sulfate concentrations up to 50% saturation, but Stv-13 had the highest solubility at 90% saturation of ammonium sulfate. Biotin binding slightly increased the solubility of the core streptavidins. Similar to the solubility changes without biotin, Stv-13 showed slightly lower solubility than Stv-25 and natural core streptavidin at ammonium sulfate concentrations up to 50% saturation but had the highest solubility at 70 and 90% saturation in the presence of biotin.

The three core streptavidins showed high solubility (greater than 75%) at ethanol concentrations up to 70% in the absence of biotin (Fig. 3C). At an ethanol concentration of 90%, only about 30% of molecules remained soluble for all of the three core streptavidins. Biotin binding had a slight effect on the solubility (Fig. 3D). There is no marked difference in the solubility characteristics in ethanol among the three core streptavidin species.

Although Stv-13 lacks two charged amino acid residues, Glu-14 and Lys-134, and two polar residues, Ser-136 and Ser-139, there is no significant difference in solubility characteristics among the three core streptavidin species. Although the relative solubilities, determined for each core streptavidin construct, may not indicate the true solubilities because the protein solutions may not have reached equilibrium under the incubation conditions used, these results clearly indicate that the terminal regions of core streptavidins have minimal effects on the solubility characteristics, unlike those of full-length streptavidin.

Structural Stability of Core Streptavidins—One characteristic that has made streptavidin one of the most frequently used proteins in various biological analyses is its extremely high structural stability. This allows, for example, conjugation of streptavidin to partner molecules by using covalent chemistry without disturbing its biotin binding ability. Very tight subunit association of streptavidin also contributes to the overall stability. For example, streptavidin remains tetrameric even in the presence of SDS or urea (33, 39, 40). The subunit association becomes even tighter upon biotin binding, and the tetrameric structure can be partly maintained by heat treatment in the presence of SDS, under which conditions, streptavidin without biotin dissociates completely into subunits (33). This
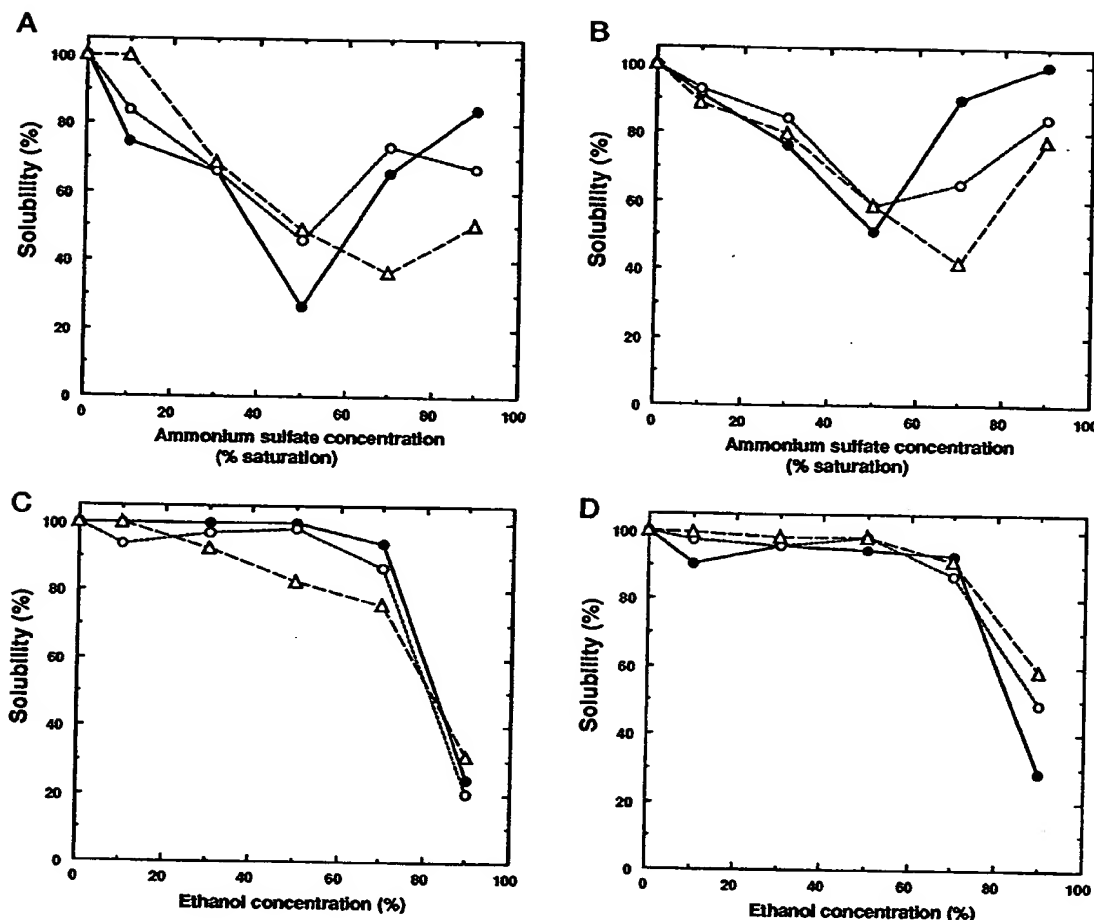
Fig. 3. **Solubility characteristics of core streptavidins in the absence and the presence of biotin.** *A,* ammonium sulfate without biotin; *B,* ammonium sulfate with biotin; *C,* ethanol without biotin; *D,* ethanol with biotin. ●, Stv-13; ○, Stv-25; □, natural core streptavidin. In the absence of biotin, each core streptavidin solution (100 μl, 5.7 nmol subunits/ml) was mixed with 1.1 ml of an appropriate ammonium sulfate or ethanol solution in Tris-buffered saline. The mixture was allowed to stand at 30 °C for 30 min and centrifuged to remove insoluble materials. The amount of soluble streptavidin in the supernatant fraction was estimated from the biotin binding ability determined by gel filtration (31). In the presence of biotin, the procedure was almost the same as above, but the biotin binding sites of each core streptavidin were saturated by adding an equimolar amount of D-[*carbonyl*-^{14}C]biotin prior to the addition of ammonium sulfate or ethanol. The amount of streptavidin in the final supernatant was estimated from the radioactivity derived from bound biotin. The fraction of original streptavidin remaining in the supernatant to the total is defined as the relative solubility, indicated in percent.

tighter subunit association upon biotin binding is essential for maintenance of bound biotin, because dissociated subunits have a much reduced biotin binding affinity due to the lack of intersubunit contacts made by Trp-120 to biotin through the dimer-dimer interface (36, 37). Extremely harsh conditions are required to effectively release bound biotin from streptavidin. Although the known three-dimensional structure of core streptavidin shows no apparent contact of the terminal residues to biotin, the disordered structure of the terminal regions (19, 20) may affect the overall stability of the protein.

To investigate how the terminal regions affect the overall stability of streptavidin, urea gradient-PAGE (41) was performed using polyacrylamide gels with a urea concentration gradient from 0 to 10 M, along with an acrylamide concentration gradient from 12 to 8%. As a control (41), urea gradient-PAGE analysis showed a marked decrease in migration of bovine serum albumin at high urea concentrations (data not shown), indicating the unfolding of the molecule. However, the three core streptavidins, Stv-13, Stv-25, and natural core

streptavidin, showed no appreciable changes in migration at urea concentrations up to 10 M, indicating the extremely high structural stability of streptavidin. This suggested that more stringent denaturation conditions are needed to allow a comparison of the stability of core streptavidins.

Thus, GdnHCl, a denaturant more potent than urea, was used. At high concentrations and very acidic pH, GdnHCl effectively denature streptavidin and release bound biotin. Biotin-binding ability was used as an estimate of the structural stability. Briefly, each core streptavidin species was incubated in solutions containing various concentrations of GdnHCl at pH 7.4 or 1.5, and then the remaining biotin binding ability was determined by gel filtration (31) (Fig. 4, *A* and *B*).

At pH 7.4, almost no changes in biotin binding ability were observed for all of the core streptavidins at GdnHCl concentrations up to 4 M (Fig. 4*A*). At 6 M GdnHCl, the biotin binding ability of Stv-25 and natural core streptavidin decreased by approximately 20%, while Stv-13 showed almost no reduction in biotin binding ability, suggesting that Stv-13 has a higher
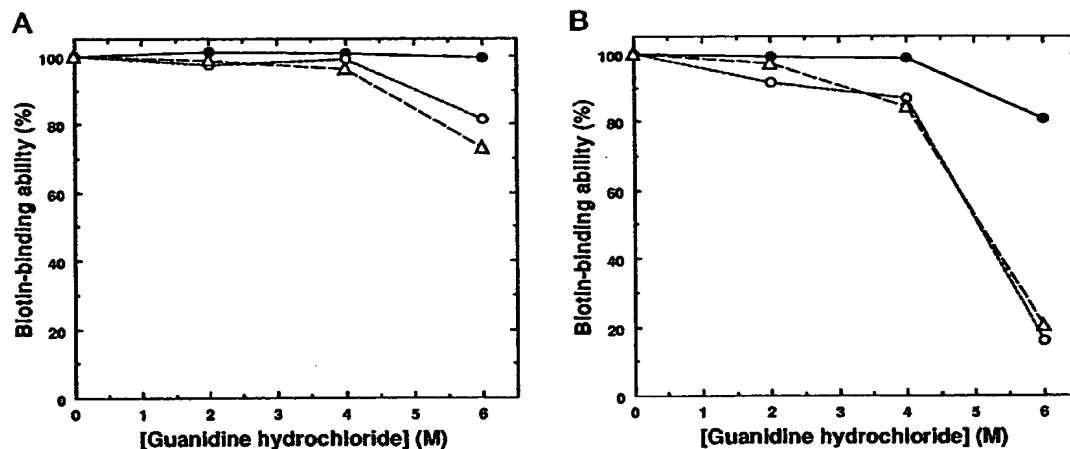
A

B

Fig. 4. **Stability of core streptavidins against denaturation by GdnHCl.** ●, Stv-13; ○, Stv-25; , natural core streptavidin. Stv-13, Stv-25, and natural core streptavidin (270 pmol of subunit/ml) were incubated at 22 °C for 10 min in GdnHCl solutions at pH 7.4 (*A*) or 1.5 (*B*). Then, the biotin binding ability of each core streptavidin was determined by gel filtration (31). All three core streptavidins bound greater than 0.96 molecules of biotin/subunit at pH 7.4 without GdnHCl, and the biotin binding ability remaining in GdnHCl is indicated in percent.

stability against denaturation by GdnHCl than Stv-25 and natural core streptavidin.

The enhanced structural stability of Stv-13 over Stv-25 and natural core streptavidin was observed even more clearly at pH 1.5 (Fig. 4*B*). Stv-13 retained almost full biotin binding ability at GdnHCl concentrations up to 4 M. In contrast, Stv-25 and natural core streptavidin lost approximately 15% of the biotin binding ability at 4 M GdnHCl. At 6 M GdnHCl, Stv-13 retained greater than 80% of the biotin binding ability, while only about 20% of the biotin binding ability was retained with both Stv-25 and natural core streptavidin.

These results demonstrate that Stv-13 has an enhanced stability against denaturation by GdnHCl when compared with Stv-25 and natural core streptavidin. This implies that the terminal regions reduce the overall structural stability of streptavidin.

Ability of Core Streptavidins to Bind Biotinylated Macromolecules—Full-length or only partially truncated streptavidin has a lower accessibility to biotinylated macromolecules than natural core streptavidins (10), because of steric hindrance caused by the terminal regions located on the surface of the molecule. To estimate how the terminal sequences of core streptavidin affect the binding to biotinylated macromolecules, the biotinylated DNA-binding ability of two core streptavidin species, Stv-13 and natural core streptavidin, was investigated. Briefly, an end-biotinylated double-stranded DNA target was mixed with core streptavidins at various ratios, and the mixtures were separated by agarose gel electrophoresis followed by staining the DNA targets with ethidium bromide.

Gel electrophoretic analysis of core streptavidin-biotinylated DNA mixtures (Fig. 5) shows that larger amounts of dimeric and trimeric biotinylated DNA targets, which are connected via single streptavidin molecules, were formed with Stv-13 than with natural core streptavidin at any molar ratio of streptavidin subunit to biotin used. Correspondingly, smaller amounts of DNA targets without streptavidin or with single streptavidin molecules (only slightly retarded from free DNA targets that were not well resolved under the electrophoresis conditions used) were observed with Stv-13. Although this analysis is not quantitative, the result indicates that Stv-13 has an enhanced binding ability for biotinylated DNA over natural core streptavidin. The enhanced binding ability of Stv-13 for biotinylated DNA reveals that the terminal regions, present on the surface

[Streptavidin subunit]

[Biotin]

M 0 Ex 0.25 0.5 0.75 1.0

— Trimeric
— Dimeric

— Monomeric
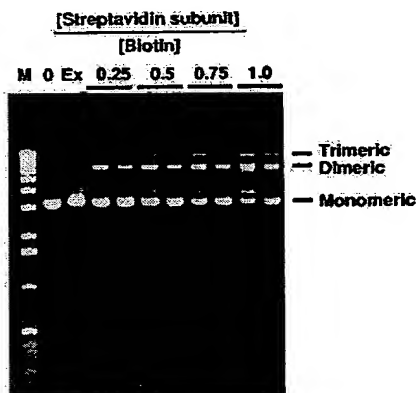
Fig. 5. **Ability of core streptavidins to bind biotinylated DNA.** A 3179-base pair end-biotinylated double-stranded DNA target was mixed with Stv-13 or natural core streptavidin at various ratios and incubated at 37 °C for 90 min. Then, the mixtures were electrophoresed on a 1.0% agarose gel and the DNA targets were stained with ethidium bromide. The *left* and *right lanes* for each molar ratio of streptavidin subunit to biotin from 0.25 to 1.0 are with Stv-13 and natural core streptavidin, respectively. The lanes marked *0* and *Ex* are with the biotinylated DNA target alone and with an excess amount of natural core streptavidin (molar ratio of streptavidin subunit to biotin ~1,000), respectively. Each lane contains 290 ng of the biotinylated DNA target. The lane marked *M* contains a 1-kilobase pair DNA ladder (Life Technologies, Inc.).

of natural core streptavidin, sterically hinder the biotin binding sites and prevent biotinylated macromolecules from approaching the biotin binding sites due, presumably, to their disordered structure.

REFERENCES

1. Chalet, L., Miller, T. W., Tausig, F., and Wolf, F. J. (1963) *Antimicrob. Agents Chemother.* **3**, 28–32
2. Chalet, L., and Wolf, F. J. (1964) *Arch. Biochem. Biophys.* **106**, 1–5
3. Green, N. M. (1975) *Adv. Protein Chem.* **29**, 85–133
4. Green, N. M. (1990) *Methods Enzymol.* **184**, 51–67
5. Wilchek, M., and Bayer, E. A. (1990) *Methods Enzymol.* **184**, 5–13
6. Wilchek, M., and Bayer, E. A. (1990) *Methods Enzymol.* **184**, 14–45

7. Argaraña, C. E., Kuntz, I. D., Birken, S., Axel, R., and Cantor, C. R. (1986) *Nucleic Acids Res.* 14, 1871–1882
8. Bayer, E. A., Ben-Hur H., Gitlin, G., and Wilchek, M. (1986) *J. Biochem. Biophys. Methods* 13, 103–112
9. Pähler, A., Hendrickson, W. A., Kolks, M. A. G., Argaraña, C. E., and Cantor, C. R. (1987) *J. Biol. Chem.* 262, 13933–13937
10. Bayer, E. A., Ben-Hur, H., Hiller, Y., and Wilchek, M. (1989) *Biochem. J.* 259, 369–376
11. Blankenburg, R., Meller, P., Ringsdorf, H., and Salesse, C. (1989) *Biochemistry* 28, 8214–8221
12. Kurtyan, J., and Weis, W. I. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 2773–2777
13. Helm, C. A., Knoll, W., and Israelachvili, J. N. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 8169–8173
14. Miyamoto, S., and Kollman, P. A. (1993) *Protein Struct. Funct. Genet.* 16, 226–245
15. Leckband, D. E., Schmitt, F.-J., Israelachvili, J. N., and Knoll, W. (1994) *Biochemistry* 33, 4611–4624
16. Florin, E.-L., Moy, V. T., and Gaub, H. E. (1994) *Science* 264, 415–417
17. Moy, V. T., Florin, E.-L., and Gaub, H. E. (1994) *Science* 266, 257–259
18. Vajda, S., Weng, Z., Rosenfeld, R., and DeLisi, C. (1994) *Biochemistry* 33, 13977–13988
19. Hendrickson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A., and Phizackerley, R. P. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 2190–2194
20. Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J., and Salemme, F. R. (1989) *Science* 243, 85–88
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

22. Sayers, J. R., Krekel, C., and Eckstein, F. (1992) *BioTechniques* 13, 592–596
23. Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol.* 189, 113–130
24. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* 185, 60–89
25. Sano, T., and Cantor, C. R. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 142–146
26. Sano, T., and Cantor, C. R. (1991) *Biochem. Biophys. Res. Commun.* 176, 571–577
27. Sano, T., Smith, C. L., and Cantor, C. R. (1993) *Bio/Technology* 11, 201–206
28. Hofmann, K., Wood, S., Brinton, C. C., Montibeller, J. A., and Finn, F. M. (1980) *Proc. Natl. Acad. Sci. U. S. A.* 77, 4666–4668
29. Sano, T., Smith, C. L., and Cantor, C. R. (1992) *Science* 258, 120–122
30. Sano, T., Glazer, A. N., and Cantor, C. R. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 1534–1538
31. Wei, R.-D. (1970) *Methods Enzymol.* 18A, 424–427
32. Laemmli, U. K. (1970) *Nature* 227, 680–685
33. Sano, T., Pandori, M. W., Smith, C. L., and Cantor, C. R. (1994) in *Advances in Biomagnetic Separation* (Uhlén, M., Hornes, E., and Olsvik, Ø., eds) pp. 21–29, Eaton Publishing, Natick, MA
34. Schmidt, T. G. M., and Skerra, A. (1994) *J. Chromatogr.* 676, 337–345
35. Thompson, L. D., and Weber, P. C. (1993) *Gene (Amst.)* 136, 243–246
36. Chilkoti, A., Tan, P. H., and Stayton, P. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 1754–1758
37. Sano, T., and Cantor, C. R. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 3180–3184
38. Sano, T., and Cantor, C. R. (1991) *Bio/Technology* 9, 1387–1391
39. Sano, T., and Cantor, C. R. (1990) *J. Biol. Chem.* 265, 3369–3373
40. Kurzban, G. P., Bayer, E. A., Wilchek, M., and Horowitz, P. M. (1991) *J. Biol. Chem.* 266, 14470–14477
41. Creighton, T. E. (1979) *J. Mol. Biol.* 129, 235–264

Evidence for multiple forms of biotin holocarboxylase synthetase in pea (*Pisum sativum*) and in *Arabidopsis thaliana*: subcellular fractionation studies and isolation of a cDNA clone

Ghislaine TISSOT, Roland DOUCE and Claude ALBAN*

Laboratoire Mixte CNRS/Rhône-Poulenc (UM41 associée au Centre National de la Recherche Scientifique), Rhône-Poulenc Agrochimie, 14–20 rue Pierre Baizet, 69263 Lyon cedex 9, France
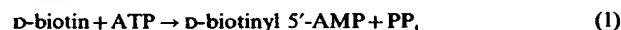
The intracellular compartmentation of biotin holocarboxylase synthetase has been investigated in pea (*Pisum sativum*) leaves, by isolation of organelles and fractionation of protoplasts. Enzyme activity was mainly located in cytosol (approx. 90% of total cellular activity). Significant activity was also identified in the soluble phase of both mitochondria and chloroplasts. Two enzyme forms were separated by anion-exchange chromatography. The major form was found to be specific for the cytosol compartment, whereas the minor form was present in mitochondria as well as in chloroplasts. We also report the isolation and DNA sequence of a cDNA encoding an *Arabidopsis thaliana* biotin holocarboxylase synthetase. This cDNA was isolated by functional complementation of a conditional lethal *Escherichia coli birA* (biotin ligase gene, which regulates biotin synthesis) mutant. This indicated that the recombinant plant protein was able to biotinylate specifically an essential apoprotein substrate in the bacterial host, that is a subunit of acetyl-CoA carboxylase called biotin carboxyl carrier protein. The full-length nucleotide sequence (1534 bp) encodes a protein of 367 amino acid residues with a molecular mass of 41172 Da and shows specific regions of similarity to other biotin holocarboxylase synthetase genes as isolated from bacteria and yeast, and with cDNA species from human. A sequence downstream of the first translation initiation site encodes a putative peptide structurally similar to organelle-targeting pre-sequences, suggesting a mitochondrial or chloroplastic localization for this isoform.

INTRODUCTION

Biotin is a small coenzyme (vitamin H or B$_8$), synthesized by plants, most bacteria and some fungi, which occurs primarily in a protein-bound state within the cell. Biotinylated proteins use this prosthetic group as a carrier of activated carboxy groups during carboxylation and decarboxylation enzymic reactions. In all organisms, these carboxylases play housekeeping functions, such as acetyl-CoA carboxylase (EC 6.4.1.2; ACCase), which catalyses the first committed step in fatty acid biosynthesis (for a general review see [1]).

Escherichia coli contains only a single biotinylated protein called biotin carboxyl carrier protein (BCCP), which functions as a subunit of ACCase. Biotinylation of apo-BCCP occurs through the action of a biotin ligase (EC 6.3.4.10). This enzyme catalyses the post-translational attachment of D-biotin to a specific Lys residue of newly synthesized apo-BCCP, via an amide linkage between the biotin carboxyl group and a unique ε-amino group of a Lys residue [2]. This covalent attachment, essential for the enzymic activation of ACCase, occurs in two distinct steps as follows:

$$\text{D-biotin} + \text{ATP} \rightarrow \text{D-biotinyl 5'-AMP} + \text{PP}_i \qquad (1)$$

$$\text{D-biotinyl 5'-AMP} + \text{apo-BCCP} \rightarrow \text{BCCP} + \text{AMP} \qquad (2)$$

Biotin ligase has been purified from *E. coli* and its gene cloned [3,4]. This enzyme, also called BirA, is a 33.5 kDa protein [4] that also acts as a repressor of the biotin operon [5]. Its three-dimensional structure has recently been determined at 2.3 Å resolution [6]. The complete nucleotide sequences of two other bacterial genes that encode proteins homologous with the *E. coli* biotin:apoprotein ligase have been reported from *Paracoccus denitrificans* [7] and *Bacillus subtilis* [8] respectively. Mammalian cells also contain biotin-dependent carboxylases that are localized in different cell compartments, i.e. ACCase in the cytosol, and 3-methylcrotonoyl-CoA carboxylase (EC 6.4.1.4), propionyl-CoA carboxylase (EC 6.4.1.3) and pyruvate carboxylase (EC 6.4.1.1) in mitochondria. As in bacteria, their activation from apo- to holo- (biotinylated) forms requires the action of a biotin ligase. Previous studies have demonstrated biotin ligase activity in both the cytosol and the mitochondria from mammalian cells [9,10], suggesting that biotinylation of biotin-dependent carboxylases occurs in their sites of enzymic activities. The corresponding enzymes from various mammalian species, which carry out the same reaction as bacterial biotin ligase, have been purified and were referred to as biotin holocarboxylase synthetase (HCS) [9,11]. Recently, clones encoding *Saccharomyces cerevisiae* HCS gene [12] and human HCS cDNA species have been obtained [13,14]. In plants, biotin and biotinylated proteins play a central role in metabolism. For example, a mutation in the biotin synthetic pathway is lethal for *Arabidopsis thaliana* [15], and plant ACCase is the target site of potent herbicides [16]. Biotin-dependent carboxylases are also present in different compartments of plant cells. As in bacteria and mammals, biotinylation of these enzymes is catalysed by HCS. Indeed, in a previous paper we provided the first direct evidence for the existence of HCS activity in plants [17]. In particular, we showed that the

---

partly purified enzyme, as obtained from pea leaves, was able to biotinylate specifically bacterial apo-BCCP as a substrate [17]. Nevertheless it is unknown whether plant biotin-dependent carboxylases are biotinylated within the cytosol and then translocated into their final site of accumulation, or are targeted into organelles as apo-proteins and subsequently biotinylated.

In this study we analysed purified chloroplasts and mitochondria from pea leaves and also used protoplasts from pea leaves as a source of cytosol with low contamination. We conclude from these experiments that, in pea leaves, HCS activity can be detected in the cytosol, mitochondria and chloroplasts.

In contrast, on the basis of the observation that plant HCS was able to biotinylate a bacterial substrate, we set out to clone a plant HCS cDNA by a functional complementation approach using an *E. coli* strain carrying a temperature-sensitive *birA* mutation and a plant cDNA expression library prepared from *A. thaliana*. This approach proved successful for isolating human HCS cDNA clones and yeast HCS gene [12,14], thus demonstrating the capacity of HCS from eukaryotes in replacing the biotin ligase function of *E. coli* BirA protein. By this method we have, in the present study, cloned and characterized for the first time a cDNA encoding HCS from the plant kingdom.

MATERIALS AND METHODS

Materials

D-[8,9-³H]Biotin (42 Ci/mmol) and [³⁵S]methionine (1000 Ci/mmol) were purchased from Amersham. D-Biotin, carbenicillin, dithiothreitol (DTT), ATP, thiamin hydroxychloride and 2,3,5-triphenyltetrazolium hydroxychloride (TTC) were obtained from Sigma Chimie SARL. Isopropyl β-D-thiogalactoside was obtained from Bioprobe Systems. Casein hydrolysate for vitamin-free assay was from Difco, and trichloroacetic acid was from Merck.

Plant material

Pea (*Pisum sativum* L., var. Douce Provence) plants were grown from seeds in soil for 8–10 days under a 12 h photoperiod of white light from fluorescent tubes (10–40 μE/s per m²) at 18 °C. The plants were watered each day with tap water.

Preparation of pea leaf crude extract

Pea leaves (2 g) were harvested and ground in liquid nitrogen with a mortar and pestle. The powder was then homogenized with 10 ml of 20 mM Tris/HCl (pH 7.8), 1 mM EDTA, 1 mM DTT, 5 mM ϵ-aminohexanoic acid and 1 mM benzamidine/HCl. The suspension was centrifuged at 72000 g for 30 min (50 Ti rotor, Beckman). The supernatant comprised the crude extract. All procedures were performed at 4 °C.

Preparation and fractionation of pea leaf protoplasts

Pea leaf protoplasts were purified from young leaves (10 days old) by the method of Baldet et al. [18]. Pea leaves (10–15 g) were cut into fine strips (1 mm) and placed in a medium containing 10 mM Mes/NaOH (pH 5.5)/0.5 M sorbitol/1 mM CaCl₂/0.05 % (w/v) PVP-25 (buffer A). After vacuum infiltration, leaf strips were washed twice in buffer A, then placed at 25 °C for 90 min in buffer A containing 2 % (w/v) cellulase Onozuka R10, 0.5 % (w/v) Macerozyme R10 and 0.2 % (w/v) Pectolyase Y-23 (Yakult Honsha Co., Shingikancho, Nishinomiya, Japan). All subsequent procedures were performed at 4 °C. Protoplasts were released from the digested tissues by gentle shaking and filtration through a 100 μm nylon mesh. The filtrate was centrifuged at

100 g for 5 min (swinging-bucket rotor). The pelleted protoplasts were resuspended in 50 ml of 0.5 M sorbitol/1 mM CaCl₂/20 mM Tris/HCl (pH 7.8) (buffer B). Two additional washes were performed in buffer B before the protoplasts were finally resuspended in buffer B supplemented with 5 mM ϵ-aminohexanoic acid and 1 mM benzamidine/HCl, at a chlorophyll concentration of 0.1 mg/ml. Protoplasts were gently ruptured by being passed first through a 20 μm nylon mesh and then through a 10 μm nylon mesh. About one-tenth of the lysed protoplast suspension was used as the protoplast crude extract. The remaining protoplast lysate was centrifuged at 300 g for 5 min. The pelleted chloroplasts were suspended in buffer B containing 5 mM ϵ-aminohexanoic acid and 1 mM benzamidine/HCl. The supernatant was then centrifuged at 12000 g for 20 min, resulting in the mitochondrial (pellet) and cytosolic (supernatant) fractions. The mitochondrial pellet was added to the chloroplastic fraction to yield the organelle fraction.

Preparation of purified chloroplasts

Young pea leaves (9 days old) were homogenized in 330 mM sorbitol/50 mM Hepes/NaOH (pH 8)/1 mM EDTA/5 mM DTT with a Waring blender. Intact chloroplasts were rapidly prepared and purified with Percoll gradients as previously described [19]. The morphological integrity of purified chloroplasts was of the order of 95 % as judged by ferricyanide reduction by untreated and osmotically shocked organelles respectively [20]. Intact chloroplasts were lysed in a buffer containing 20 mM Tris/HCl (pH 7.8), 1 mM EDTA, 1 mM DTT, 5 mM ϵ-aminohexanoic acid and 1 mM benzamidine/HCl at a final protein concentration of 40 mg/ml after one freeze-thawing cycle (5 min in liquid nitrogen, thawing at 25 °C). The suspension of broken chloroplasts was centrifuged at 72000 g over a 0.6 M sucrose cushion for 20 min. The pellet and the supernatant comprised the chloroplast membranes (envelope membranes and thylakoids) and the soluble fraction (stroma) respectively. All procedures were performed at 4 °C.

Preparation of purified mitochondria

Mitochondria were isolated and purified from young pea leaves by using self-generating Percoll gradients as described by Douce et al. [21]. The morphological integrity of purified mitochondria was greater than 95 % as determined by the rate of KCN-sensitive cytochrome c-dependent O_2 uptake in untreated and osmotically shocked mitochondria respectively [21]. Total lysis of mitochondria was achieved by three freeze–thawing cycles (5 min in liquid nitrogen, thawing at 25 °C), performed in the same buffer as that used for the lysis of purified chloroplasts (see above), at a protein concentration of 40 mg/ml. The suspension of broken mitochondria was centrifuged at 72000 g over a 0.6 M sucrose cushion for 20 min. The pellet and the supernatant comprised the mitochondrial membranes and the soluble fraction (matrix) respectively. All procedures were performed at 4 °C.

Measurement of marker enzyme activities

Except where otherwise noted, enzymes were assayed spectrophotometrically at 340 nm by coupling to oxido-reduction of NADH (or NADPH) in 1 ml reaction volumes. Triton X-100 (0.05 %, w/v) was added to each reaction mixture. The activity of each enzyme in pea leaf extract was strictly dependent on the presence of all necessary substrates and cofactors, and linear with respect to the amount of extract assayed.

Glyceraldehyde-3-phosphate dehydrogenase

Glyceraldehyde-3-phosphate dehydrogenase (NADP$^+$-dependent) (EC 1.2.1.13; GraPDH) activity was measured by the conversion of 1,3-bisphosphoglycerate into glyceraldehyde 3-phosphate [22]. The assay mixture contained 50 mM Hepes/NaOH, pH 7.8, 1 mM MgCl$_2$, 4 mM EDTA, 5 mM DTT, 0.2 mM NADPH, 1 mM ATP, 10 units of phosphoglycerate kinase (EC 2.7.2.3), 10 units of triose-phosphate isomerase (EC 5.3.1.1) and 10 mM 3-phosphoglycerate.

Fumarase

Activity was determined by the method of Hill and Bradshaw [23], by following the appearance of fumarate spectrophotometrically at 250 nm from reaction assays containing 50 mM Hepes/NaOH, pH 8, and 50 mM malate.

Pyrophosphate:fructose-6-phosphate 1-phosphotransferase

Pyrophosphate:fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90; PFP) activity was measured as described by Weiner et al. [22]. The assay mixture contained 50 mM Tricine/NaOH, pH 7.8, 0.5 mM MgCl$_2$, 10 μM fructose 2,6-bisphosphate, 5 mM fructose 6-phosphate, 10 units of triose-phosphate isomerase, 1 unit of glycerol-3-phosphate dehydrogenase (EC 1.1.1.8), 0.1 unit of aldolase (EC 4.1.2.13), 0.15 mM NADH and 0.6 mM Na$_4$P$_2$O$_7$.

Thermolysin treatment of purified intact organelles

Proteolytic digestion was performed in accordance with a previously published procedure [24]. Intact chloroplasts or mitochondria (final concentration 10 mg/ml protein) were incubated for 1 h at 4 °C in the following medium: 0.3 M sucrose/10 mM Tricine/NaOH (pH 7.8)/1 mM CaCl$_2$/0.2 mg/ml thermolysin from *Bacillus thermoproteolyticus* (Boehringer). Thermolysin is active only in the presence of Ca^{2+} ions and its activity can be easily inhibited by the addition of 10 mM EGTA. Polypeptides localized in the inner envelope membrane or in the stroma of chloroplasts, and also in the inner membrane or the matrix space of mitochondria, are not hydrolysed during the incubation because thermolysin is unable to cross the outer envelope membrane of chloroplasts or the outer membrane of mitochondria. In addition, the integrity of organelles is maintained during incubations under mild conditions (at 4 °C, with low thermolysin concentration) [24]. Therefore after incubation for 1 h in the presence of thermolysin under these conditions, chloroplasts and mitochondria were re-purified on a Percoll gradient containing protease inhibitors (1 mM PMSF, 1 mM benzamidine/HCl and 5 mM ε-aminohexanoic acid) to remove the protease and the broken organelles. The treated, intact organelles were recovered, stored on ice and assayed for HCS activity; the soluble enzymes were then analysed by chromatography on a Mono Q anion-exchange column.

Mono Q anion-exchange chromatography

All chromatography experiments were performed at 4 °C with Mono Q HR 5/5 (Pharmacia) coupled to an FPLC system (Pharmacia) to obtain precise and repeatable elution patterns. Soluble protein extracts (crude leaf extract, chloroplast stroma and mitochondrial matrix) prepared as described above were desalted on a Sephadex G-25 (M) column in a medium containing 20 mM Tris/HCl, pH 8, 1 mM EDTA, 1 mM DTT, 1 mM benzamidine/HCl and 5 mM ε-aminohexanoic acid and then loaded (2–5 mg) on the anion-exchange column pre-equilibrated

with the same buffer. After loading, the column was washed with 5 ml of buffer and eluted with the following linear NaCl gradients: 0–0.3 M NaCl (30 ml), 0.3–0.5 M NaCl (10 ml), 1 M NaCl (5 ml) in the same buffer at a flow rate of 0.5 ml/min. Fractions of 1 ml were collected and assayed for HCS activity.

Latency measurements

Corrections were made for extra-organellar activity by comparing the activities in ruptured and intact organelles. The organelles were kept intact by adding an osmoticum to the reaction medium (0.3 M sucrose) or ruptured by adding 0.05 % (w/v) Triton X-100. The percentage of latent activity is the ratio of organellar (ruptured minus intact) activity to the total (ruptured) activity. We verified that, under these conditions, enzymic activities were not affected by the presence of the detergent.

Protein and chlorophyll determinations

Protein was measured by the method of Bradford [25] using Bio-Rad protein assay reagent (Bio-Rad Laboratories) with bovine γ-globulin as a standard. Chlorophyll was measured by the method of Arnon [26].

Bacterial strains and growth conditions

A temperature-sensitive *E. coli birA215* mutant (strain BM4050) lacking HCS activity *in vitro* was generously provided by Dr. A. M. Campbell (Stanford University, Stanford, CA, U.S.A.) [27,28]. Mutations in the *birA* gene affect the biotin ligase function of the BirA protein, resulting in biotin auxotrophy. This strain grows normally at 30 °C on minimal medium M9 (48 mM Na$_2$HPO$_4$/22 mM KH$_2$PO$_4$/19 mM NH$_4$Cl/8.5 mM NaCl/1 mM MgSO$_4$/0.1 mM CaCl$_2$) supplemented with 0.2 % glucose, 2 nM biotin, 0.4 % casein hydrolysate, 100 μg/ml TTC and 1 μg/ml thiamin. In contrast, bacteria fail to grow under the same conditions (either in liquid medium or on plate) but at a temperature of 43 °C, since at this temperature the mutant biotin:apoprotein ligase has a greatly decreased affinity for biotin. When required, carbenicillin was added at 100 μg/ml. Competent cells, grown in Luria–Bertani broth, were prepared according to the method of Dower et al. [29].

Measurement of HCS activity

HCS activity was determined by measuring the covalent attachment of D-[^{3}H]biotin to bacterial apo-BCCP by a trichloroacetic acid precipitation assay as described previously [17]. The apo-BCCP-protein-substrate was prepared from the *E. coli birA215* mutant grown on glucose minimal medium supplemented with limiting amounts of D-biotin (0.4 nM) [17]. The reaction mixture contained the following components in a total volume of 200 μl: 5 mM ATP, 5 mM MgCl$_2$, 0.2 mM DTT in 20 mM phosphate buffer (pH 7.5), 10–200 μg of enzyme extract, 180 μg of apo-BCCP extract and 250 nM D-[^{3}H]biotin. The reaction was initiated by addition of the labelled D-biotin. Incubations were for 20–60 min at 37 °C. Then 125 μl aliquots of the reaction mixture were collected on glass microfibre filters (Whatman GF/C) and proteins were precipitated by five washes of the filters in 10 % (w/v) trichloroacetic acid. The filters were washed once with ethanol and dried; radioactivity was counted in scintillation vials containing 8 ml of scintillation liquid (Ready protein; Beckman). Duplicate assays without apo-BCCP and/or HCS extract were run as controls.

Isolation of A. thaliana cDNA clones encoding HCS by functional complementation of E. coli birA215

A λYES-R cDNA expression library prepared with mRNA species from A. thaliana plants was first converted into a plasmid library carrying ampicillin resistance, as described by Elledge et al. [30]. E. coli birA215 competent cells were transformed by electroporation with this plasmid library by using a Bio-Rad Gene Pulser operating at 15 kV/cm pulse, 25 μF and 200 Ω. Then bacteria were suspended in 1 ml of SOC broth [31] and incubated at 37 °C with shaking (250 rev./min) for 1 h. Cells were washed twice in M9 medium to eliminate free biotin and plated on M9 plates, supplemented with 2 nM biotin, 0.4% casein hydrolysate, 100 μg/ml TTC, 1 μg/ml thiamin and 100 μg/ml carbenicillin. Cell culture was conducted at 43 °C to select rescued clones. Transformation efficiency was determined by cell growth at 30 °C on identical M9 plates. Eight successive electroporations were performed as described above, with a transformation frequency of approx. 3×10^6 transformants per μg of plasmid DNA. Plasmids were isolated [31] from survivor colonies growing at 43 °C on M9/carbenicillin plates, and the original E. coli birA215 mutant was transformed again with the purified plasmids to confirm the functional complementation. Only transformants that survived the second round of selection were used for further analyses.

Preparation of bacterial crude extract

To assess the recombinant HCS activity, the biotin auxotroph bacterial strain E. coli birA215 (strain BM4050) (untransformed and transformed) as well as the control wild-type strain E. coli BirA+ (strain BM2661) were grown on Luria–Bertani liquid medium (supplemented with carbenicillin in the case of transformed bacteria) at 37 °C, until A_{600} reached 0.5. After induction of the pYES lac promoter with 1 mM isopropyl β-D-thiogalactoside for 4 h, cells were collected by centrifugation at 3000 g for 15 min (JA 20 rotor; Beckman) and suspended in buffer A containing 20 mM phosphate buffer, pH 7.5, 1 mM Na_2EDTA, 1 mM DTT and a mixture of protease inhibitors (1 mM PMSF, 5 mM ε-aminohexanoic acid and 1 mM benzamidine/HCl). Cells were then disrupted by sonication at 0 °C and lysates were centrifuged for 15 min at 15000 g to remove cell debris (JA 20 rotor; Beckman). The supernatant was desalted by passage through a Sephadex G25 (M) column (Pharmacia) equilibrated in buffer A.

Subcloning and DNA sequencing

The cDNA inserts were subcloned into the EcoRI site of vector plasmid pBluescript II SK(−) (Stratagene). DNA sequence analysis was performed on both strands by using Prism Kit with fluorescent dideoxynucleotides, Taq DNA polymerase (Applied Biosystems) and T3 and T7 universal primers. In addition, specific oligonucleotide primers were used for further sequencing. Gene Works 2.4 and PCGENE (IntelliGenetics) software packages were used for sequence analyses.

Transcription–translation of pBluescript HCS in vitro

Transcription–translation in vitro of pBluescript HCS (which contains the entire region of A. thaliana HCS cDNA cloned under the control of the T7 promoter) was done with T7 RNA polymerase and a rabbit reticulocyte lysate from Novagen kit (Single Tube Protein® System 2, T7) and [35S]methionine (Amersham) in accordance with the instructions of the manufacturer. For the reaction the DNA template (2 μg of pBluescript

HCS-2) was transcribed at 30 °C for 15 min, followed by translation reaction (in the presence of 40 μCi of [35S]methionine) at 30 °C for 60 min. A control reaction lacking a DNA template was performed under the same conditions. The translational products were precipitated with 10% (w/v) trichloroacetic acid. After a 1 h incubation at 0 °C and centrifugation at 30000 g for 15 min at 4 °C (JA 20 rotor; Beckman), the precipitated protein pellets were washed with cold acetone, air-dried and resuspended in 50 mM Tris/HCl (pH 8)/1 mM PMSF/5 mM ε-aminohexanoic acid/1 mM benzamidine/HCl. Polypeptides labelled with [35S]methionine were analysed by SDS/PAGE and fluorography. SDS/PAGE [10% (w/v) gel] was performed at room temperature in slab gels (15 cm × 15 cm). The experimental conditions for gel preparation, sample solubilization, electrophoresis and gel staining were as detailed by Chua [32]. After staining, gels were soaked in Amplify solution (Amersham) and dried before fluorography for 1 week on X-ray-films.

Genomic Southern blot analysis

Total DNA was isolated from young A. thaliana plants. Approx. 2–5 μg of DNA was digested overnight with appropriate restriction enzymes (New England Biolabs) and fragments were separated by 0.8% agarose gel electrophoresis. After blotting to a Hybond-N+ membrane (Amersham), hybridization was performed with the entire A. thaliana HCS-2 cDNA or the EcoRI–EcoRV 5′ end of the pBluescript HCS-2 cDNA probes that had been 32P-labelled with a random priming kit (Pharmacia) [31]. The nucleic acid hybridization solution was composed of 6 × SSC (the stock solution was 20 × SSC, containing 175.3 g/l NaCl and 88.2 g/l sodium citrate, pH 7), 0.5% SDS and 0.25% low-fat dried milk. Hybridization proceeded overnight at 65 °C and membranes were washed at 65 °C for 30 min in 1 × SSC/0.1% SDS and 0.1 × SSC/0.1% SDS.

RESULTS

Intracellular localization of HCS activity in pea leaves

Protoplasts, isolated by enzymic digestion from young pea leaves (9 days old), were fractionated by gently rupturing through a fine nylon mesh followed by centrifugation to yield an organelle fraction (pellet) and a cytosolic fraction (supernatant). These two fractions were assayed for HCS activity, and their purity was assessed by measurement of selected subcellular marker enzyme activities. Marker enzymes used were fumarase for mitochondria, GraPDH for chloroplasts, and PFP for cytosol. As shown in Table 1, most of the chloroplast (83%) and mitochondrial (87%) marker activities were recovered in the pellet. Because it was difficult to rupture the small protoplasts quantitatively without affecting chloroplast integrity, we performed an incomplete rupture of the protoplasts to obtain only minimal chloroplastic contamination in the cytosolic fraction (supernatant). Thus approx. 31% of the total cytosolic marker activity was recovered in the pellet. The supernatant obtained after differential centrifugation of lysed pea leaf protoplasts contained a substantial proportion of cytosolic marker enzymes (63% of the total activity) and was only slightly contaminated by chloroplast and mitochondrial enzymes (Table 1). HCS activity was detected both in the supernatant (59%) and in the pellet (38%). As the activity of HCS in the organelle pellet was not very different from the cytosolic contamination level, it was difficult to conclude that chloroplasts and/or mitochondria contained an HCS form. From these results it is clear that, in higher plant cells, HCS activity is located mostly in the cytosol (probably up to 90% of total cellular activity).

Table 1 Subcellular localization of HCS in pea leaf protoplasts

HCS activity was assayed in broken pea leaf protoplasts (protoplast extract, 12.5 mg of protein), organelles (pellet) and cytosol (supernatant) as described in the Materials and methods section in the presence of 0.05% (w/v) Triton X-100. Also indicated are the activities of marker enzymes to check for cross-contamination. The marker enzymes were as follows: cytosol, PFP; chloroplasts, GraPDH; mitochondria, fumarase. The results presented for distribution in the supernatant and pellet and for recovery are expressed as a percentage of the total activity recovered, and are means ± S.D. for four different experimental determinations.

	Enzyme activity			
	PFP	GraPDH	Fumarase	HCS
Activity in protoplast extract (mol/min)	154×10^{-9}	1108×10^{-9}	45×10^{-9}	0.61×10^{-12}
Distribution in pellet (%)	31 ± 5	83 ± 1	87 ± 3	38 ± 1
Distribution in supernatant (%)	63 ± 8	7 ± 2	3 ± 1	59 ± 9
Recovery (%)	94 ± 13	90 ± 3	90 ± 4	97 ± 10

Table 2 Distribution of activities of HCS and marker enzymes in Percoll-purified chloroplasts and mitochondria from pea leaves

Preparation of purified organelles, measurement of different enzymic activities in the presence of 0.05% (w/v) Triton X-100, and latent activity definition were as described in the Materials and methods section. The marker enzymes were identical with those used for the fractionation of purified pea protoplasts (Table 1). Results are from a representative experiment repeated three times. Abbreviation: n.d., not detectable.

	Enzyme activity (mol/min per mg of protein)				HCS latency (%)
Sample	PFP	GraPDH	Fumarase	HCS	
Chloroplasts	n.d.	342×10^{-9}	8×10^{-9}	0.026×10^{-12}	80
Mitochondria	n.d.	12×10^{-9}	870×10^{-9}	0.041×10^{-12}	92

With the aim of determining the possible occurrence of HCS activity in plant cell organelles, we conducted a large-scale purification of intact pea leaf chloroplasts and mitochondria on Percoll gradients. Assays of various selected marker enzymes for chloroplasts (GraPDH), mitochondria (fumarase) and cytosol (PFP) showed that this purification procedure completely eliminated cytosolic contamination. Furthermore, both the chloroplast and the mitochondrial fractions were essentially free from mitochondrial and chloroplast contamination, respectively (Table 2). HCS activity was found to be associated with both mitochondria and chloroplasts, with a specific activity of 0.041 and 0.026 pmol/min per mg of protein respectively (Table 2). On a protein basis, these values were one-quarter to one-sixth of those measured in the cytosolic fraction prepared from fractionated protoplasts (0.17 pmol/min per mg of protein). Nevertheless the total absence of cytosolic contamination and the high latency values of HCS activity measured in both purified chloroplasts and mitochondria demonstrated that the enzyme activity was present within the organelles (Table 2).

Separation of HCS activities by Mono Q anion-exchange chromatography

To determine whether HCS activities found in pea leaves represented different forms of the enzyme, extracts from pea leaves (crude leaf extract, soluble proteins from Percoll-purified chloroplasts and mitochondria, and cytosol from leaf proto-
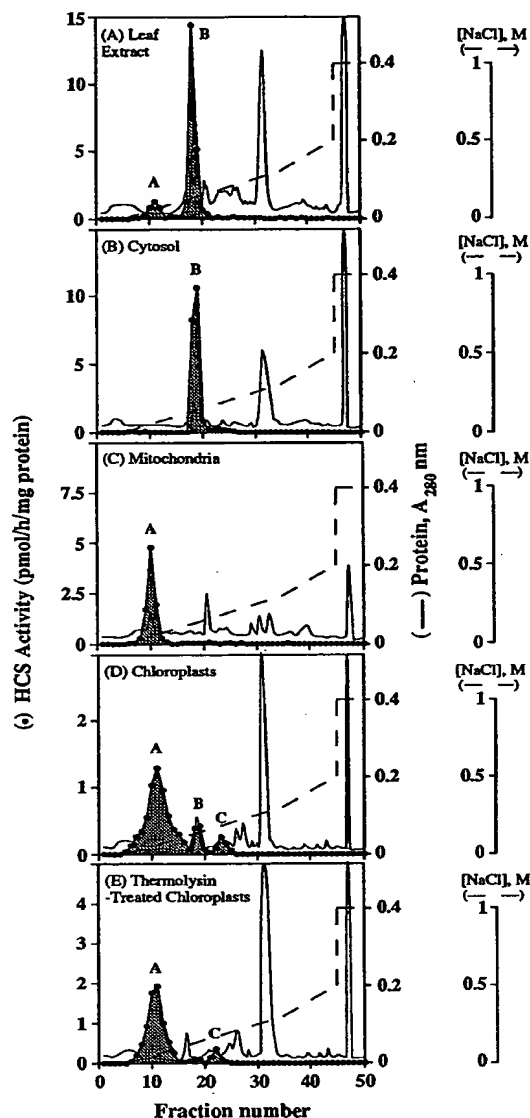
Figure 1 Separation of multiple forms of HCS from pea leaves by ion-exchange chromatography

Pea leaf extracts were fractionated by chromatography on a Mono Q HR5/5 column in 20 mM Tris/HCl (pH 8)/1 mM EDTA/1 mM DTT/5 mM ε-aminohexanoic acid/1 mM benzamidine/HCl at 4 °C. After being loaded, the column was washed with 5 ml of buffer and eluted with a NaCl gradient at 0.5 ml/min. Fractions of 1 ml were collected and assayed for HCS activity. The samples loaded were: (A) leaf crude extract (5 mg); (B) cytosol (2 mg); (C) mitochondrial extract (2.3 mg); (D) chloroplast extract (4.9 mg); (E) extract from thermolysin-treated chloroplasts (3.8 mg).

plasts) were fractionated on a Mono Q HR 5/5 column. Recovery of HCS activity was in the range 85–90% for each of the chromatographies performed. The elution profile obtained for the leaf crude extract (Figure 1A) shows that HCS activity can be resolved into two peaks (A and B). The major peak (peak B),

eluted at 140 mM NaCl, represented approx. 90% of the total activity. Peak A, eluted at 50–60 mM NaCl, contained approx. 10% of the total activity present in the crude extract. The cytosolic pattern (Figure 1B) exhibited only one peak of HCS activity, displaying chromatographic properties similar to those of peak B in the crude extract. In purified mitochondria (Figure 1C), HCS activity was also eluted as a single peak, but with a profile similar to that of peak A in the crude extract. Finally, the elution profile obtained with the chloroplast extract (Figure 1D) contained three peaks of HCS activity. The major peak (85% of the total activity) and one of the minor peaks (10% of the total activity) had similar profiles to those of peak A and peak B in the crude extract respectively. The last minor peak detected in chloroplast extracts was eluted at approx. 170 mM NaCl, and was referred to as peak C. From these results we can conclude that peak A is associated with both mitochondria and chloroplasts, and peak B with cytosol. Peak C, which was detected in chloroplasts but not in crude extracts, might correspond to a degraded form of chloroplast HCS. Finally, peak B, the component of the cytosolic fraction, was always found in the chloroplast extract, although cytosolic contamination of chloroplasts was always negligible (Table 2). This observation led us to question whether some cytosolic HCS could be loosely adsorbed to the outer surface of the outer membrane of the chloroplast envelope, to be released during the breakage of chloroplasts by osmotic shock and one cycle of freeze–thawing. To verify this hypothesis, we used a mild proteolytic digestion of intact chloroplasts with thermolysin because this non-penetrating proteolytic enzyme has been demonstrated to be an efficient tool for characterizing those envelope proteins that are accessible from the cytosolic side of the outer membrane [24]. The elution pattern of the extract obtained from thermolysin-treated chloroplasts confirms that HCS activity was indeed a genuine constituent of chloroplasts because peaks A and C were not affected by the treatment (Figure 1E). In contrast, peak B was no longer detectable in thermolysin-treated chloroplasts, thus providing clear evidence for an extra-plastidial localization of this HCS activity. Therefore, from the results presented in Figures 1(B), 1(D) and 1(E), and in Tables 1 and 2, we can conclude that peak B does indeed correspond to HCS present in the cytosol. As a control we used the same proteolysis treatment to confirm that HCS activity associated with the mitochondria was clearly inside the organelle. The identical elution profiles before and after thermolysin treatment confirmed the high latency of the activity associated with the mitochondria.

Isolation of cDNA clones encoding HCS by complementation of temperature-sensitive BM4050 cells

To characterize the HCS isoforms further at the molecular level, we developed a functional complementation screening technique, using a higher plant cDNA expression library, and an E. coli mutant affected in HCS activity. The E. coli birA215 competent cells lacking endogenous biotin ligase activity when grown at 43 °C were transformed by electroporation with approx. 6 × 10⁸ plasmids expressing an A. thaliana cDNA library (initially containing 10⁷ independent recombinants [30]). Isolation of E. coli birA215-complemented clones was attempted by selection on M9 plates containing 0.2% glucose, 2 nM biotin, 0.4% casein hydrolysate, 100 µg/ml TTC, 1 µg/ml thiamin hydroxychloride and 100 µg/ml carbenicillin at 43 °C. After 48 h of growth at 43 °C, six colonies were isolated. When these clones were cultured at 43 °C in M9 liquid medium supplemented as above, four of them retained the ability to complement the biotin auxotrophy

Figure 2   Nucleotide and predicted amino acid sequences of the cDNA encoding HCS from A. thaliana (HCS-2 clone)

The coding sequence is indicated with capital letters; the non-coding sequence is indicated with lower-case letters. Nucleotides are numbered at the right. The first in-frame ATG is in bold and the corresponding stop codon is marked with an asterisk. The longest open reading frame extends for 1101 bp and translates into a 367-residue protein with a molecular mass of 41 172 Da.

and temperature-sensitive growth of the birA215 host, and were referred to as HCS-1, HCS-2, HCS-3 and HCS-4.

Characterization, nucleotide sequences and deduced amino acid sequences of A. thaliana HCS cDNA

The length of the cDNA inserts was found to be approx. 1.4 kb in HCS-1 and HCS-4, and 1.5 kb in HCS-2 and HCS-3. The four cDNA inserts shared the same internal sequence, but differed only in the length of their untranslated 5′ and 3′ ends. Thus HCS-2 and HCS-3 were found to be identical, whereas HCS-1 carried a 33 bp extension on the 5′ end and HCS-4 had a 23 bp deletion on its 5′ end compared with HCS-2 and HCS-3. Finally, all four clones contained a poly(A) tail, although none of them exhibited a typical eukaryotic polyadenylation signal sequence [33]. The position of the poly(A) tail with respect to the TGA stop codon varied depending on the clone. Thus the distance between the poly(A) tail and the TGA stop codon was 315 bp for clones HCS-2 and HCS-3, and 252 bp for clones HCS-1 and HCS-4.

The complete nucleotide sequence of the longest isolated cDNA (HCS-2) is shown in Figure 2 along with the deduced

Figure 3   HCS activity in crude protein extracts from E. coli birA215 strain transformed or not with the four HCS cDNA clones and from E. coli birA+ wild-type strain

Total soluble proteins from the different strains were extracted after induction of the Lac promoter with isopropyl β-D-thiogalactoside, for an optimal time of 4 h, as described in the Materials and methods section. HCS activity was determined as described previously by measuring the specific incorporation of D-[³H]biotin into bacterial apo-BCCP [17]. D-Biotin (1 pmol) corresponds to 95000 d.p.m.

amino acid sequence. HCS-2 is a 1534 bp cDNA including one large open reading frame of 1101 bp. The first in-frame ATG occurring at nucleotide 50 of the cDNA initiates the longest open reading frame present on this cDNA and encodes a predicted polypeptide of 367 residues with a molecular mass of 41172 Da. A second in-frame ATG is present 114 bp downstream from the first one. The nucleotide sequence around the first ATG codon, TTTA<u>ATG</u>GA (positions 46–54), differs from the plant consensus translation initiation motif, AACA<u>A</u>UGGC [34]. In contrast, that for the second, AGCA<u>ATG</u>GA (positions 160–168), more closely matches the plant consensus sequence. Finally, the presence of an in-frame nonsense codon TGA, located 27 bp upstream from the first ATG, confirms that HCS-2 is full length.

Functional characterization of HCS cDNA

To confirm that the four cDNA species encoded HCS, we measured this enzyme's activity in crude protein extracts obtained from the E. coli birA215 mutant complemented with each of the four clones, by using bacterial apo-BCCP as the biotin acceptor substrate (Figure 3). All four clones had the same orientation in pYES vector and were under the control of the lac promoter. Whereas no HCS activity could be detected with the birA215 mutant, the four complemented clones exhibited significant HCS activity (Figure 3). Thus levels of HCS activity from the recombinant clones were 13–14-fold higher than that from the wild-type birA+ (BM2661) strain (Figure 3).

Transcription–translation of pBluescript HCS-2 in vitro

Because the complete nucleotide sequence contains two in-frame ATG sequences, it is possible that isolated HCS cDNA species encode two distinct polypeptides. Coupled transcription–translation of HCS-2 cDNA in vitro subcloned in pBluescript under the control of the T7 promoter, using T7 RNA polymerase and a rabbit reticulocyte lysate, produced two major translation products of approx. 37 and 41 kDa, i.e. of the expected sizes for an initiation of translation at the two in-frame ATG sequences
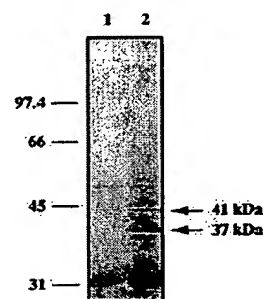
Figure 4   Coupled transcription–translation of pBluescript HCS-2 in vitro

Polypeptides radioactively labelled with [³⁵S]methionine were subjected to SDS/PAGE and analysed by fluorography as described in the Materials and methods section. Lane 1, fluorography of radiolabelled polypeptides obtained from a control reaction lacking DNA template; lane 2, fluorography of translation products obtained with 2 μg of pBluescript HCS-2. The positions of molecular mass markers (given in kDa) are indicated on the left.

(Figure 4, lane 2). The additional labelled polypeptide of molecular mass 52 kDa did not correspond to a translation product of the HCS-2 cDNA because it was also detected in the control experiment lacking the DNA template (Figure 4, lane 1). It seems unlikely that the production of the two polypeptides can be explained by a premature termination of translation because a similar result was obtained with wheat germ extract as the translation system in vitro (results not shown).

Southern blot analysis

Southern analysis was used to examine the number of genes encoding HCS in A. thaliana. Total DNA was digested with restriction enzymes that cut once (EcoRV) or do not cut (EcoRI, HindIII) within the cDNA, and the fragments were resolved by agarose gel electrophoresis. After transfer to a Hybond-N+ membrane, the resultant blot was probed with the ³²P-labelled complete HCS-2 insert as described in the Materials and methods section. As shown in Figure 5, digestion with EcoRI or HindIII produced two bands (7.1–4.6 and 8–3.4 kb), whereas digestion with EcoRV resulted in three hybridization bands of 13, 12 (arrowheads) and 1.7 kb. After double digestion of total DNA with EcoRV and HindIII, the probe detected three bands, a 1.7 kb band generated by EcoRV, a 3.4 kb band generated by HindIII and a 6.2 kb band. Hybridizations were also performed with a restriction fragment corresponding to the 501 bp 5' end (EcoRI–EcoRV fragment) of the pBluescript HCS-2 cDNA. After digestion with EcoRI or HindIII a similar pattern of hybridization was obtained, whereas after digestion with EcoRV or double digestion with EcoRV and HindIII the 5'-end probe revealed only two bands in both cases, that is the 13 and 1.7 kb bands and the 3.4 and 1.7 kb bands respectively (results not shown). Altogether, these results are consistent with the existence of two related genes encoding HCS in A. thaliana. However, such a banding pattern could also be generated by a single gene containing large introns. This last possibility cannot be completely ruled out.

DISCUSSION

Fractionation of pea leaf protoplasts and purification of chloroplasts and mitochondria from this tissue clearly indicate that HCS activity is associated with several subcellular compartments.
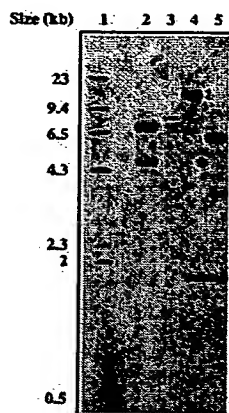
Figure 5   Southern blot hybridization of total A. thaliana DNA with the complete HCS cDNA as a probe

Total leaf DNA from A. thaliana was digested (2–5 μg per reaction) with EcoRI (lane 2), HindIII (lane 3), EcoRV (lane 4) and HindIII–EcoRV (lane 5). Lane 1 corresponds to size markers (λDNA HindIII digest). DNA restriction fragments were separated on a 0.8% agarose gel, transferred to Hybond-N+ membrane and then hybridized with the ^{32}P-labelled probe corresponding to the pBluescript HCS-2 cDNA excised with EcoRI.

HCS was resolved into two peaks by anion-exchange chromatography. The main peak (making up approx. 90% of the total activity) was located in the cytosol. The other peak was present in both chloroplasts and mitochondria. The great purity and the higher latency values of HCS activity measured in Percoll-purified chloroplasts and mitochondria, together with the protection of the enzyme activity in these organelles during thermolysin treatment, confirmed that HCS is a genuine constituent of chloroplasts and mitochondria.

Detailed characterization of the incorporation of D-biotin into plant biotin-dependent carboxylases requires large quantities of HCS isoenzymes. To overcome the very low abundance of these enzymes in plants, as judged by the low specific activities detected in each cell compartment, we attempted to obtain the cDNA species encoding these isoforms in order to over-express it. The biotinylation reaction is evolutionarily conserved such that biotin ligases function with various apocarboxylases across species boundaries [2]. We have therefore used the cross-species reactivity of biotin ligases to clone a cDNA encoding a complete A. thaliana HCS by functional complementation of a mutant birA strain of E. coli. Selection of plant cDNA species encoding HCS by this technique confirms our previous results showing that pea HCS efficiently biotinylates bacterial apo-BCCP in vitro [17].

Comparison of the predicted protein sequence of the isolated plant HCS cDNA with those compiled in the GenBank and EMBL databases showed low but significant similarity to biotin: apoprotein ligases from E. coli (17% identity; 33% similarity) [4], B. subtilis (18% identity; 31% similarity) [8], P. denitrificans (14% identity; 25% similarity) [7], Homo sapiens (24% identity; 40% similarity) [13,14] and S. cerevisiae (22% identity; 40% similarity) [12] (Figure 6). Specific areas of similarity are restricted to a region known in E. coli BirA to contain the biotin-binding site [6]. Thus, across a restricted 125-residue region (positions 122–246 of A. thaliana HCS), the plant enzyme shares 33% identity (51% similarity) with BirA protein from E. coli, 33% identity (47% similarity) with a putative homologue of BirA protein from B. subtilis, 30% identity (47% similarity) with the

candidate for BirA protein from P. denitrificans, 30% identity (49% similarity) with HCS from H. sapiens and 34% identity (51% similarity) with HCS from S. cerevisiae. Most importantly, the eight residues involved in direct contact with biotin in BirA from E. coli, as determined by X-ray crystallography [6], are strictly conserved in A. thaliana HCS. These residues are Ser[122], Thr[123], Gln[145], Arg[149], Lys[220], Gly[223], Gly[240] and Gly[242]. Interestingly, similarity to chicken avidin, the protein with the highest known affinity for biotin [35] was also observed, particularly in the region located between Asp[205] and Thr[229] of the predicted amino acid sequence (Figure 6). The sequence identity in this zone was 36%, and increased to 60% when conservative substitutions were included. Altogether these results confirm the finding that this region is essential for biotin binding. Within this region, the sequence GRGRTK is present at positions 148–153 and partly matches that found in other biotin ligases (GRGRRG) (Figure 6). Although no crystallographic evidence for the ATP-binding site has been demonstrated [6], the structure GXGXXG has been associated with ATP binding in several enzymes [36,37]. However, it also seems to be involved in contact with biotin in BirA [6]. Presumably this reflects a requirement for ATP and biotin to be spatially close to permit the formation of biotinyl 5'-adenylate, an intermediate of the biotinylation reaction. Sequence comparisons within the helix–turn–helix DNA binding motif of E. coli and B. subtilis BirA proteins (residues 22–46 and 23–47 respectively) [4,8] revealed no region of similarity, suggesting that in contrast with what occurs in bacteria, plant HCS is not involved in the repression of biotin synthesis. This observation is in good agreement with previous findings showing that the free D-biotin concentration in the cytosol of plant cells is of the order of 11 μM [17,18]. Indeed, this level, which is approx. 2000-fold that found in bacteria [38], was not compatible with the existence of a strong repressor of biotin synthesis in plants, comparable with the repressor function of BirA protein, which regulates the level of biotin in bacteria. Finally, there are sequences conserved between A. thaliana HCS and the two other eukaryotic proteins that are not found in BirA from bacteria (Figure 6). For example, in the C-terminal portion of the predicted A. thaliana HCS sequence (residues 301–367), the plant enzyme shares only 12% identity (19% similarity) with BirA from B. subtilis compared with 37% identity (58% similarity) with HCS from S. cerevisiae.

Comparison of the N-terminal portion of known HCS sequences indicates that the A. thaliana protein contains an N-terminal extension of approx. 30 amino acids compared with those from bacteria (Figure 6). This N-terminal extension is rich in hydrophobic, hydroxylated and positively charged amino acids, but poor in acidic residues. In addition, theoretical secondary-structure predictions indicate that this sequence would fold into an amphiphilic α-helix. Thus this extension has many characteristics in common with mitochondrial or chloroplast presequences. Furthermore, analysis of the primary sequence of the plant HCS sequence by the subroutine TRANSEP from the program PCGENE (IntelliGenetics) predicted that this protein might be targeted to an organelle, possibly the mitochondrion, and identified the sequence Arg[29]-Leu/Ser-Phe as a putative cleavage-site motif [39]. However, the occurrence of an in-frame Met residue at position 39, which might act as a potential translation initiation site, as indicated by transcription–translation experiments in vitro (Figure 4), might indicate that the present clone encodes, in reality, a cytosolic HCS isoform. Therefore it is not possible from our data to determine the exact subcellular localization of the cloned A. thaliana HCS. Further studies on the uptake of the cloned plant HCS by mitochondria and/or chloroplasts, as well as in situ localization experiments on plant cells overexpressing this clone are in progress to make a

compartments, namely a mitochondrial 3-methylcrotonoyl-CoA carboxylase, a chloroplastic ACCase (of eukaryotic type in Gramineae and of prokaryotic type in other plants) and a putative cytosolic ACCase [40–45]. Thus the existence of HCS isoforms in different cell compartments suggests that the different biotin-dependent carboxylases in plants are biotinylated in the cell compartment within which they are localized. In a previous publication we showed that pea leaf cells contain a pool of free D-biotin localized in the cytosol, and no detectable levels of this vitamin in the organelles [18]. This therefore raises the question of how chloroplast and mitochondrial HCS forms might function *in vivo*. More recently we demonstrated that plant HCS displayed a very low K_m for D-biotin of 28 nM [17]. As the detection limit for the biological assay used to determine free D-biotin levels in pea leaf cell compartments was of the order of 0.05 ng [18], it is possible that chloroplasts and mitochondria actually contain very low D-biotin concentrations that are sufficient to allow the biotinylation of biotin-dependent carboxylases present in these organelles by HCS isoenzymes. Further characterization of these enzymes, and particularly the identification of a possible structurally distinct HCS isoform, will be necessary for understanding the mechanism of biotinylation of apocarboxylases in plants and to elucidate the question of why HCS activity is compartmentalized in plant cells.

Figure 6 Amino acid sequence comparison between biotin ligases and chicken avidin

Sequences are shown for *A. thaliana* HCS, *E. coli* BirA protein [4], *B. subtilis* BirA protein [8], *P. denitrificans* putative biotin ligase [7], *H. sapiens* HCS [13,14], *S. cerevisiae* HCS [12] and chicken avidin [35]. The sequences are aligned with gaps (–) to maximize identity. Conserved amino acids represented in bold are reported in the deduced consensus line (Consensus). Conservative amino acid substitutions were determined in accordance with the following grouping: I-L-M-V, N-Q, R-K-H, A-S-P-T-G, Y-F-W and D-E. The amino acid positions for *H. sapiens* and *S. cerevisiae* HCS and chicken avidin are shown at the right.

definitive assignment of the cellular localization of this clone to a specific compartment. It is interesting to note a recent proposal that human cytosolic and mitochondrial HCS isoforms are synthesized from a single species of mRNA either by alternative translational initiation at two in-frame AUGs or by a splicing mechanism [14].

Finally, the presence of HCS isoforms in different cell compartments, as determined biochemically in pea leaves, raises the question of their physiological significance in plants. In mammals, two isoforms of the same HCS targeted to cytosol and mitochondria respectively have been characterized, catalysing the biotinylation of one cytosolic biotin-dependent carboxylase (ACCase) for the former and three mitochondrial biotin-dependent carboxylases (3-methylcrotonoyl-CoA carboxylase, propionyl-CoA carboxylase and pyruvate carboxylase) for the latter [9–11,14]. In plant cells, we and others have characterized different biotin-dependent carboxylases localized in three cell

REFERENCES

1 Dakshinamurti, K. and Chauhan, J. (1989) Vitam. Horm. **45**, 337–384

2 Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C. and Wood, H. G. (1988) J. Biol. Chem. **263**, 6461–6464

3 Eisenberg, M. A., Prahash, O. and Hsiung, S.-C. (1982) J. Biol. Chem. **257**, 15167–15173

4 Howard, P. K., Shaw, J. and Otsuka, A. J. (1985) Gene **35**, 321–331

5 Cronan, Jr., J. E. (1989) Cell **58**, 427–429

6 Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. and Matthews, B. W. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 9257–9261

7 Xu, X., Mastsuno-Yagi, A. and Yagi, T. (1993) Biochemistry **32**, 968–981

8 Bower, S., Perkins, J., Yocum, R. R., Serror, P., Sorokin, A., Rahaim, P., Howitt, C. L., Prasad, N., Ehrlich, S. D. and Pero, J. (1995) J. Bacteriol. **177**, 2572–2575

9 Chiba, Y., Suzuki, Y., Aoki, Y., Ishida, Y. and Narisawa, K. (1994) Arch. Biochem. Biophys. **313**, 8–14

10 Chang, H. I. and Cohen, N. D. (1983) Arch. Biochem. Biophys. **225**, 237–247

11 Xia, W.-L., Zhang, J. and Ahmad, F. (1994) Biochem. Mol. Biol. Int. **34**, 225–232

12 Cronan, Jr., J. E. and Wallace, J. C. (1995) FEMS **130**, 221–230

13 Suzuki, Y., Aoki, Y., Ishida, Y., Chiba, Y., Iwamatsu, A., Kishino, Y., Niikawa, N., Matsubara, Y. and Narisawa, K. (1994) Nat. Genet. **8**, 122–128

14 León-Del-Rio, A., Leclerc, D., Akerman, B., Wakamatsu, N. and Gravel, R. A. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 4626–4630

15 Shellhammer, J. and Meinke, D. (1990) Plant Physiol. **93**, 1162–1167

16 Harwood, J. L. (1988) Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**, 101–138

17 Tissot, G., Job, D., Douce, R. and Alban, C. (1996) Biochem. J. **314**, 391–395

18 Baldet, P., Alban, C., Axiotis, S. and Douce, R. (1993) Arch. Biochem. Biophys. **303**, 67–73

19 Mourioux, G. and Douce, R. (1981) Plant Physiol. **67**, 470–473

20 Walker, D. A., Cerovic, Z. G. and Robinson, S. P. (1987) Methods Enzymol. **148**, 145–157

21 Douce, R., Bourguignon, J., Brouquisse, R. and Neuburger, M. (1987) Methods Enzymol. **148**, 403–417

22 Weiner, H., Stitt, M. and Heldt, H. W. (1987) Biochim. Biophys. Acta **893**, 13–21

23 Hill, R. L. and Bradshaw, R. A. (1969) Methods Enzymol. **13**, 91–99

24 Dorne, A. J., Block, M. A., Joyard, J. and Douce, R. (1982) FEBS Lett. **145**, 30–34

25 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254

26 Arnon, D. I. (1949) Plant Physiol. **24**, 1–15

27   Barker, D. F. and Campbell, A. M. (1981) J. Mol. Biol. **146**, 451–467

28   Barker, D. F. and Campbell, A. M. (1981) J. Mol. Biol. **146**, 469–492

29   Dower, W. J., Miller, J. F. and Ragsdale, C. W. (1988) Nucleic Acids Res. **16**, 6127–6145

30   Elledge, S. J., Mulligan, J. T., Ramer, S. W., Spottswood, M. and Davis, R. W. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 1731–1735

31   Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

32   Chua, N.-H. (1980) Methods Enzymol. **69**, 434–436

33   Joshi, C. P. (1987) Nucleic Acids Res. **15**, 9627–9640

34   Lückte, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. and Scheele, G. A. (1987) EMBO J. **6**, 43–48

35   Gope, M. L., Keinanen, R. A., Kristo, P. A., Conneely, O. M., Beattie, W. G., Zarucki-Schutz, T., O'Malley, B. W. and Kulomaa, M. S. (1987) Nucleic Acids Res. **15**, 3595–3606

36   Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) EMBO J. **1**, 945–951

37   Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W. and Hermodson, M. A. (1986) Nature (London) **323**, 448–450

38   Eisenberg, M. A. (1973) Adv. Enzymol. **38**, 317–372

39   Von Heinje, G. (1992) Genet. Eng. **14**, 1–11

40   Baldet, P., Alban, C., Axiotis, S. and Douce, R. (1992) Plant Physiol. **99**, 450–455

41   Alban, C., Baldet, P., Axiotis, S. and Douce, R. (1993) Plant Physiol. **102**, 957–965

42   Egli, M. A., Gengenbach, B. G., Gronwald, J. W., Somers, D. A. and Wyse, D. L. (1993) Plant Physiol. **101**, 499–506

43   Alban, C., Baldet, P. and Douce, R. (1994) Biochem. J. **300**, 557–565

44   Konishi, T. and Sasaki, Y. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 3598–3601

45   Choi, J. K., Yu, F., Wurtele, E. S. and Nikolau, B. J. (1995) Plant Physiol. **109**, 619–625

Role of avidin and other biotin-binding proteins in the deposition and distribution of biotin in chicken eggs

Discovery of a new biotin-binding protein

Harold B. WHITE III* and Colin C. WHITEHEAD
AFRC Poultry Research Centre, Roslin, Midlothian EH25 9PS, Scotland, U.K.

In addition to the previously characterized egg-yolk biotin-binding protein (BBP-I), we have discovered another BBP (BBP-II) in the plasma and yolk from laying hens. BBP-I is stable to 65 °C, whereas BBP-II is stable to 45 °C. Both proteins are normally saturated with biotin and together they account for most, if not all, of the biotin in hen plasma and yolk, except in hens fed excessive amounts of biotin (> 1 mg of biotin/kg of feed). The maximal production of BBP-I is attained at lower levels of dietary biotin (~ 50 μg/kg) than for BBP-II (~ 250 μg/kg); however, the maximal production of BBP-II is severalfold greater than for BBP-I. Consequently, as dietary biotin increases, the ratio of BBP-II to BBP-I increases and becomes constant at dietary intakes of biotin above 250 μg/kg. The observation that the amounts of these proteins are limited by biotin in the normal dietary range (< 250 μg/kg) suggests that biotin is required for the synthesis, secretion or stability of these proteins. Although both plasma vitamin–protein complexes are transported to the oocyte and concentrated in the yolk, BBP-II is transferred more efficiently. Thus biotin deposition in the yolk is a function of the amounts and relative concentrations of the two proteins. Dietary biotin above 250 μg/kg exceeds the transport capacity of BBP-I and BBP-II in the plasma; however, unbound biotin does not accumulate. Rather it is efficiently scavenged by avidin in the oviduct and transferred to the egg albumen. Only when avidin becomes saturated at high dietary intake does free or weakly bound biotin accumulate in plasma and yolk. The synthesis of avidin is independent of dietary biotin. Small amounts of BBPs with the heat-stability of avidin or BBP-I respectively are present in the plasma of adult males or immature chickens. BBP-II, the major BBP in the plasma and yolk of laying hens, was not detected in the plasma of non-laying chickens.

INTRODUCTION

In 1927 Boas showed that there was a nutritional factor in egg yolk and other foods that was inactivated by a heat-labile component of egg albumen. Biotin, the nutritional factor, was first isolated and characterized from duck egg yolk (Kögel & Tönnis, 1936). Avidin, the antivitamin, was purified from egg albumen (Pennington et al., 1942) and shown to be an extraordinarily stable tetrameric protein whose affinity for biotin may well be the strongest non-covalent interaction between a protein and a small molecule (Green, 1975). György & Rose (1942) observed that biotin in egg yolk became dialysable on heating, but it was more than 30 years before a biotin-binding protein (BBP) was identified and purified from egg yolk (White et al., 1976; Meslar et al., 1978; Murty & Adiga, 1984). Though similar in size and quaternary structure, BBP is distinct from avidin by a number of criteria. This protein is also present in the plasma of laying hens (Mandella et al., 1978) and is presumed to be deposited in the ovarian follicle along with other egg-yolk proteins (White, 1985).

The original assay for yolk BBP was based on the equilibrium exchange of endogenous bound biotin and exogenous [¹⁴C]biotin at 65 °C (Meslar & White, 1979). Although this assay is quite satisfactory for partially purified BBP, it was technically difficult to perform on yolk extracts, and the limits of detection were approached with plasma samples. Furthermore, the calculations for this assay included corrections for isotope dilution which assumed that the protein was initially saturated with unlabelled biotin and that no additional endogenous biotin was present. Recently a new and much more sensitive assay based on the exchange binding of [³H]biotin has been developed (White & McGahan, 1986; H. B. White, T. McGahan & M. A. Letavic, unpublished work). The graphical analysis of this assay yields the amount of endogenous ligand (Lotter et al., 1982). Although there are still some technical problems with this assay that cause underestimates of the amount of binding proteins, it is now possible to analyse BBP in samples not accessible with the previous assay.

In the process of applying this new assay to yolk and plasma samples from laying hens that had been fed diets differing in their biotin content, we discovered that the endogenous biotin of content of yolk and plasma exceeded the BBP binding sites by severalfold. This was unexpected, because this excess biotin should have been dialysable. The paradox was resolved by the discovery of a second BBP (BBP-II) that binds most, if not all, of the biotin not bound by the previously recognized BBP hereafter designated 'BBP-I'.

Abbreviation used: BBP, biotin-binding protein.

* Present address and address for correspondence and reprint requests: Department of Chemistry, University of Delaware, Newark, DE 19716, U.S.A.

Table 1. Composition of experimental diets

Vitamin/mineral supplements provided (per kg of diet): A: copper, 3.6 mg; iodide, 0.4 mg; iron, 80 mg; magnesium, 300 mg; manganese, 100 mg; zinc, 50 mg; retinol, 600 μg; cholecalciferol, 15 μg; α-tocopherol, 17 mg; menadione, 1.3 mg; riboflavin, 4 mg; nicotinic acid, 28 mg; panthothenic acid, 10 mg; B: as for A, plus cyanocobalamin, 25 μg; folic acid, 0.5 mg; pyridoxine, 4 mg; thiamin, 2 mg; C: as for A, plus biotin, 100 μg; choline chloride, 0.5 g.

	Composition (g/kg)		
Ingredient	Laying diet	Biotin-deficient laying diet	Chick diet
Maize	—	—	50
Wheat	750	553	596
Starch	–	80	30
Herring meal	60	—	145
Meat and bone meal	30	—	—
Soybean meal	25	—	105
Casein (low-vitamin)	—	100	—
Gelatin	—	48	—
Egg albumen (spray-dried)	—	36	—
Isolated soy protein	20	—	40
Vegetable oil	—	30	—
Cellulose	—	36	—
Limestone flour	65	76	13
Dicalcium phosphate	22	30	12
Salt	3	4	3
DL-Methionine	—	2	—
Vitamin/mineral supplements			
A	5	—	—
B	—	5	—
C	—	—	3

In the present paper we document the presence of BBP-II and show that it is unstable under the conditions used for the assay of BBP-I. Furthermore we show that neither the deposition of biotin in yolk nor the distribution of biotin between yolk and albumen is a simple function of dietary biotin or plasma biotin concentration. The patterns can be explained by the differential synthesis and differential transport of BBP-I and BBP-II.

MATERIALS AND METHODS

Experimental design

ISA Brown hens (96 in all, housed in individual battery cages) were maintained on a standard laying diet for several weeks until a high rate of egg production was established. The diet, the composition of which is given in Table 1, was based on wheat and contained a relatively low amount of available biotin ($\sim$ 28 mg/kg), but was nevertheless thought to be aequate in all nutrients for maximal egg production (Whitehead, 1980). Seven groups of 12 hens each were fed the standard diet supplemented with 0, 100, 250, 500, 1000, 2000 and 4000 μg biotin/kg of feed. An eighth group was fed a biotin-deficient diet (Table 1). This diet contained hen's-egg albumen as a source of avidin and was thought to be virtually devoid of available biotin. Daily egg productions and weekly feed consumptions were recorded. Plasma, yolk and albumen were collected weekly and analysed for biotin and biotin-binding proteins.

In a second experiment, newly hatched chicks were sexed and then fed a diet (Table 1) containing about 160 mg of available biotin/kg. Blood samples were taken from several chicks and pooled at various ages and analysed for BBPs.

Sample preparation

Plasma, yolk and albumen samples were obtained exactly as described by White et al. (1986) for similar studies on riboflavin, except that the 4-fold dilutions of egg yolk were made with 50 mM-sodium acetate, pH 5.5, containing 50 mM-NaCl. Samples were stored frozen at $-20\ °C$ until assayed for BBP. On the day of analysis, plasma samples were usually diluted 10-fold, yolk samples an additional 10–50-fold and albumen samples 200-fold with the above buffer. Samples for biotin analysis were not diluted before freezing.

Assays for BBP-I

BBP-I was assayed by a radioligand-exchange procedure analogous to that described for assaying riboflavin-binding protein (Lotter et al., 1982). A series of tubes containing 0.1 μCi of D-[8,9-³H(n)]biotin (lot 2169-146, 35.0 Ci/mmol; New England Nuclear Corp., Boston, MA, U.S.A.) and 20–240 μl of diluted plasma or yolk in the above sodium acetate buffer (total volume 1.0 ml) was incubated at 65 °C for 40 min to equilibrate free and BBP-I-bound biotin. These conditions denature BBP-II. The cooled incubation mixtures were quantitatively transferred to small phosphocellulose columns [0.25 ml bed volume in polypropylene pipette tips (Sarsted, no. 91-787)]. Uncomplexed biotin was eluted with two 1.0 ml buffer washes. BBP-I-bound biotin was then eluted directly into scintillation vials by washing the columns with 4×0.25 ml of the sodium acetate buffer

containing 2 M-NaCl. A portion (10 ml) of scintillation-counting fluid (Opti Phase 'X', Amersham International) was added and the amount of radioactivity determined in a liquid-scintillation counter. Non-specific binding was determined in the presence of 500-fold excess of unlabelled biotin, and an avidin solution was used to determine total bindable radioactivity. Data were plotted according to the following linear equation (White & McGahan, 1986) (the slope and intercepts were determined by linear-regression analysis):

$$\frac{{}^{*}L_{\mathrm{T}}}{{}^{*}L_{\mathrm{B}}} = \frac{1}{[P_{\mathrm{T}}]} \cdot \frac{{}^{*}L_{\mathrm{T}} \cdot F}{V} + \frac{L_{\mathrm{T}}}{P_{\mathrm{T}}}$$

If ${}^{*}L_{\mathrm{T}}/{}^{*}L_{\mathrm{B}}$, the ratio of total to bound radioactive ligand is plotted as a function of ${}^{*}L_{\mathrm{T}} \cdot F/V$, where F and V are the dilution factor and volume of the diluted sample used in the assay; the y-intercept, $L_{\mathrm{T}}/P_{\mathrm{T}}$, is the ratio of endogenous biotin to biotin-binding sites, and the reciprocal of the slope, $[P_{\mathrm{T}}]$, is the concentration of binding protein in the undiluted sample.

Assays for BBP-II

An assay that directly measures BBP-II in the presence of BBP-I has not been developed. Two procedures have been used to estimate BBP-II activity. The first method was exactly like that described for BBP-I except that the incubation temperature was 45 °C instead of 65 °C. The choice of temperatures was based on the thermal-stability profiles presented in Figs. 5 and 6 (below). The difference between the results from the 45 °C and 65 °C assays was attributed to BBP-II, but this approximation is in error to the extent that BBP-I does not achieve equilibrium exchange at 45 °C. Alternatively BBP-II activity has been estimated from the BBP-I assays by assuming that BBP-II is saturated with biotin and all endogenous biotin in excess of BBP-I was bound to BBP-II. Although both of these assays are qualitatively reliable, the values obtained are systematically low compared with the values expected from bioassay for biotin.

Assays for avidin

Avidin was measured in albumen or plasma exactly as for BBP-I, except the incubation was at 85 °C, a temperature that destroys BBP-I and BBP-II. Unoccupied biotin-binding sites were determined at room temperature, where exchange does not occur.

Biotin analyses

Undiluted yolk, plasma and albumen samples were frozen and sent with feed samples to F. Hoffmann–LaRoche, Basel, Switzerland, where they were assayed for biotin by using Lactobacillus plantarum as described by Frigg & Brubacher (1976).

RESULTS

Biotin content of diets

The diets were formulated to contain 30, 100, 250, 500, 1000, 2000, and 4000 μg of biotin/kg. The analyses of these diets showed respectively 93, 165, 332, 622, 1054, 1831 and 3540 μg total biotin/kg. The data plotted in the Figures correspond to available biotin which is about 65 μg/kg less than the measured values. This correction is justified by the fact that over 95% of the biotin in wheat is not available (Frigg, 1976; Whitehead et al., 1982).
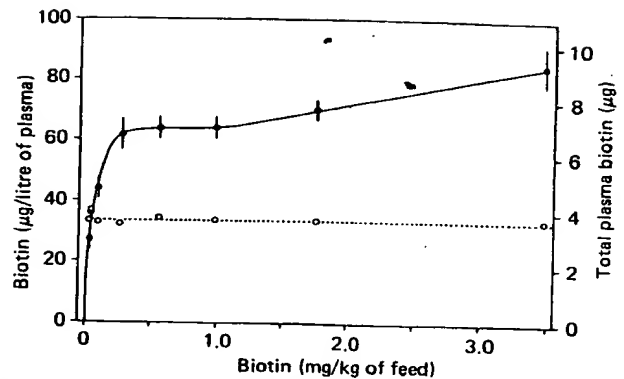

Fig. 1. **Effect of dietary biotin on the concentration of biotin the plasma of laying hens**

Open circles (O) are the control values for the treatment groups immediately before the experimental diets were begun. The dotted line represents the average of these control values (33.9 ± 1.99 μg/l). Closed circles (●) represent the average ± s.d. for samples taken after 1, 2, and 5 weeks on the experimental diets. The right-hand scale is based on a calculated blood volume of 109 ml.
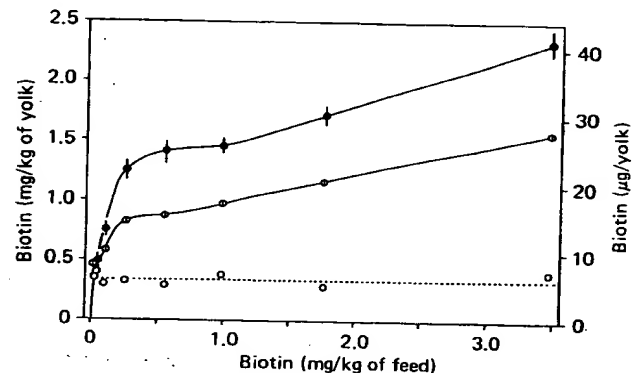

Fig. 2. **Effect of dietary biotin on the concentration of biotin in chicken egg yolk**

Open circles (O) represent control values for yolk samples from the treatment group immediately before the experimental diets were begun. The dotted line represents the average of these control values (343.1 ± 48.3 μg/kg). The divided circles (⊕) represent the average for duplicate yolk samples after the birds had been 1 week on the experimental diets. Closed circles (●) represent the mean ± s.d. for duplicate samples taken after the birds had been 2 and 5 weeks on the experimental diet. The right-hand scale is based on an average yolk weight of 17.6 g.

Biotin content of plasma, yolk and albumen as a function of dietary biotin

Figs. 1–3 show respectively the biotin levels in plasma, yolk and albumen before and after 1, 2 and 5 weeks on the experimental diets. Steady-state levels were achieved within 1 week in plasma and albumen and by 2 weeks in yolk. With respect to dietary biotin, the accumulation of biotin in plasma, yolk and albumen can be considered in three phases. In the normal range of dietary biotin (< 250 μg/kg) there is a very strong relationship
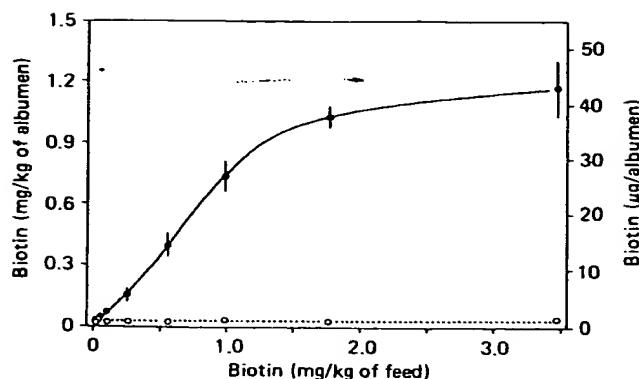
Fig. 3. Effect of dietary biotin on the amount of biotin in chicken egg albumen

Open circles (○) represent the control values for the treatment groups immediately before the experimental diets were begun. The dotted line represents the average of these control values (26.4 ± 6.7 μg/kg). Closed circles (●) represent the average $\pm$ s.D. for duplicate samples taken after the birds had been 1, 2 and 5 weeks on the experimental diets. The right-hand scale is based on a recovered albumen weight of 36.6 g/egg.
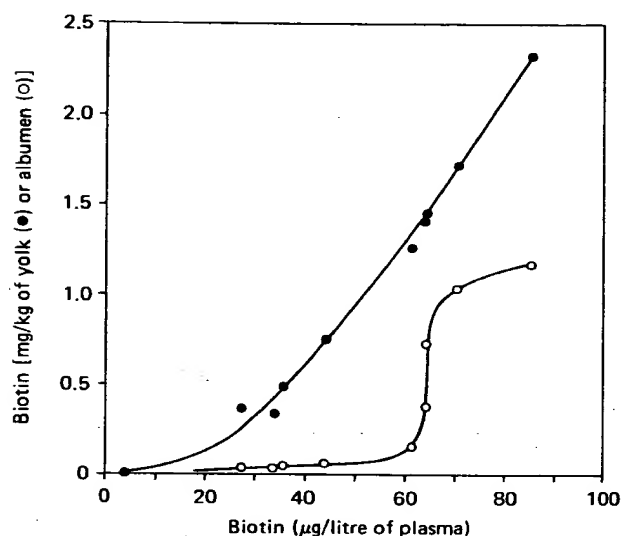
Fig. 4. Concentration of biotin in the yolk (●) and albumen (○) of chicken eggs as a function of the concentration of biotin in the plasma of the laying hen

Each point represents the average of four to six biotin determinations made on plasma, yolk and albumen samples. Note the vertical scale is 25 times that of the horizontal scale.

between dietary biotin and the biotin content of plasma and yolk, whereas the biotin content of albumen is rather low. Between 250 and 1000 μg/kg there is a plateau in the biotin contents of plasma and yolk and a large increase in the biotin content of albumen. Above 1000 μg of biotin/kg there is a further gradual increase in biotin in all three compartments. Considering the range from 30 to 3475 μg/kg, an increase of over 100-fold in dietary biotin, there is a 2.5-fold increase in plasma biotin, a

6.8-fold increase in yolk biotin and a 44.5-fold increase in albumen biotin. Hens fed the biotin-deficient diets for 6 weeks continued to lay eggs. The plasma and yolk concentrations of biotin from these birds were 4.05 μg/l and 13.0 μg/kg respectively. The biotin content of albumen was below the detection limits of the assay. These responses are broadly in agreement with the observations of Frigg et al. (1984).

Biotin content of yolk and albumen as a function of plasma biotin

The plasma distributes absorbed dietary biotin to the various tissues of the body, and thus biotin in the plasma is an intermediate between ingested biotin and biotin deposited in the yolk and albumen of eggs. Fig. 4 shows that biotin deposition in yolk and albumen is not a simple function of plasma biotin concentration. Although biotin is concentrated in yolk relative to plasma over the entire range of experimental conditions, the efficiency of transfer is greatly reduced when plasma biotin concentrations are below 20 μg/l. The yolk-to-plasma concentration ratio in this region is less than 10:1, whereas the incremental increase above this region is near 35:1. The relationship between biotin in albumen and plasma shows an even more dramatic discontinuity. Below 64 μg/l, very little biotin is deposited in albumen. A slight increase in plasma biotin above this level results in very large increases in the deposition of biotin in albumen.

BBP-I and BBP-II content of plasma and yolk

Fig. 5 shows that hens transferred from a diet containing 30 μg of biotin/kg of feed to one containing 987 μg/kg increase the production of a heat-labile BBP appearing in both the plasma and yolk. The data in Fig. 5 are not corrected for isotope dilution. If such a correction were made, there would be little difference in the total biotin binding above 60 °C, whereas difference in biotin-binding near 45 °C would be accentuated.

The kinetics of biotin exchange at 45 and 65 °C in the same yolk samples are shown in Fig. 6. Again, the increase in the amount of heat-labile BBP-II is evident in the samples from hens fed high-biotin diets. There is very little difference in the amount of BBP-I in these two samples.

Similar analyses conducted at 45 and 65 °C on yolk samples from hens fed a wide range of dietary biotin (Fig. 7) show that maximal production of BBP-I occurs on diets containing 50 μg or more of biotin per kg of feed, whereas maximal production of BBP-II occurs with higher levels of dietary biotin ($>$ 250 μg/kg). The slightly lower amounts of BBP-I at higher dietary biotin levels are not considered signficant at this time because the large amounts of endogenous biotin render the BBP-I assay less accurate in this region.

Fig. 8 shows that there is an 8–10-fold excess of endogenous biotin over BBP-I in yolk from hens fed very high levels of biotin. The mean $\pm$ s.D. of this ratio for 20 assays performed on yolk samples from pre-experimental control hens maintained on 30 μg of biotin/kg of diet is 1.64 ± 0.20. If one assumes that the excess biotin in the samples is due to BBP-II bound biotin, the ratio of BBP-II to BBP-I increases from 0.64 at 30 μg of biotin/kg to 5 at 500 μg/kg. The biotin content of yolk is included in Fig. 8 as a reference to show that the data obtained by microbiological assay are qualitatively similar to those
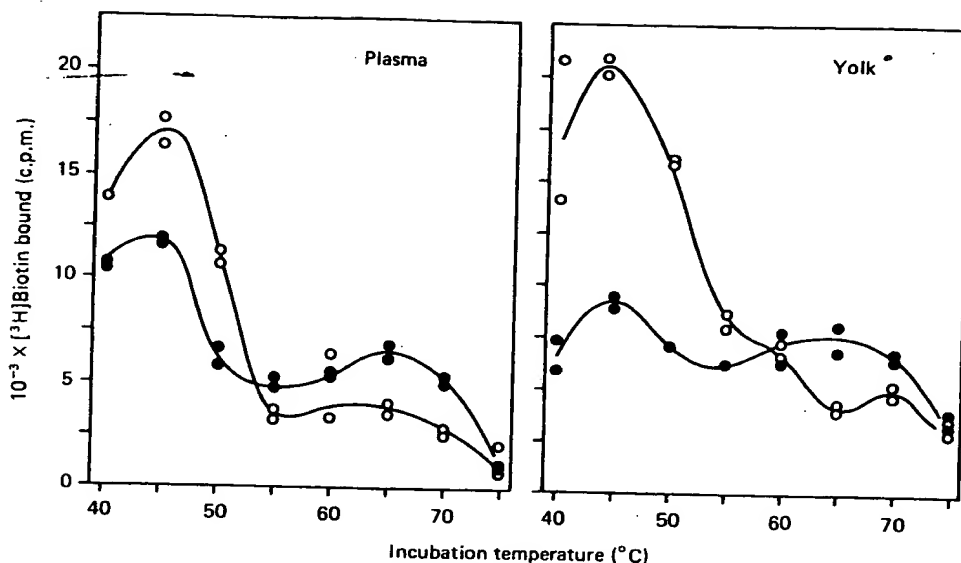
Fig. 5. Concentration of a heat-labile BBP increases in the plasma and egg yolk from hens fed diets with high concentrations of biotin

The equivalent of 10 μl of plasma and 1 μl of yolk from hens fed 28 μg of biotin/kg of feed (●) were incubated for 30 min in the presence of [³H]biotin, and the bound [³H]biotin determined. Similar samples from the same birds were analysed after they had been maintained for 3 weeks on the standard diet supplemented to 987 μg of biotin/kg (O). The amount of endogenous biotin, and thus the isotope dilution, in these samples is greater than those from the samples from hens on the low-biotin diet.

obtained by isotope-exchange assays of the binding protein. Assays at 45 °C show that BBP-II is saturated with biotin even in yolk samples obtained from hens fed a biotin-deficient diet for 5 weeks.

Content and fractional saturation of avidin in albumen

Table 2 shows that the concentration of avidin in albumen is unaffected by dietary biotin. The fractional saturation is strongly dependent on dietary biotin. Avidin is over 90% saturated with biotin when dietary biotin exceeds about 1500 μg/kg.

BBPs in the plasma of immature and adult chickens of both sexes

As noted in the Methods and materials section, the amounts of the various binding proteins cannot be measured with high accuracy when they are present in mixtures. Table 3 presents the analysis of BBP-I, BBP-II and avidin in the plasma of immature and adult male and female chickens. The identification is based on heat-stability. BBP-I is present in the plasma of immature chickens of both sexes at concentrations about one-tenth that found in laying hens. BBP-II was detected only in the plasma of laying hens. Avidin was present in
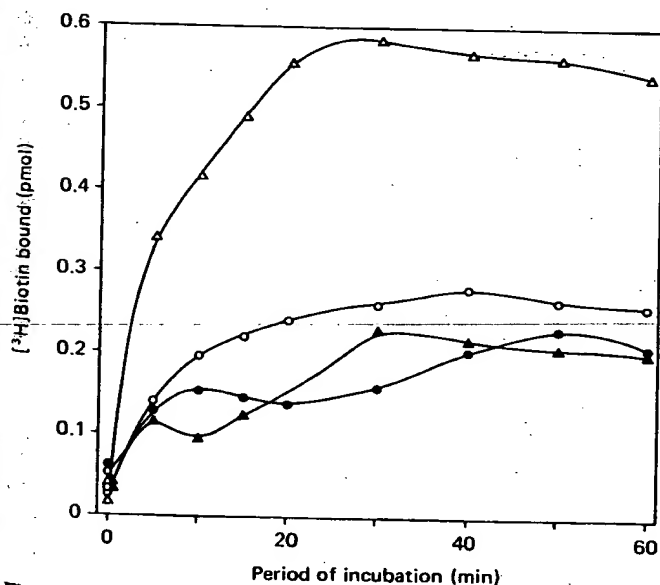

Fig. 6. Kinetics of biotin exchange at 45 °C (O, △) and 65 °C (●, ▲) show the accumulation of a heat-labile biotin-binding protein in egg yolks from hens fed a diet with 30 μg of biotin/kg of feed (O, ●) and then fed a diet with 1000 μg/kg for 3 weeks (△, ▲)
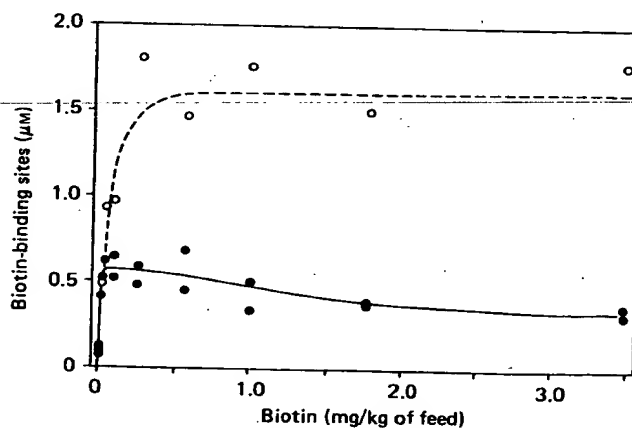

Fig. 7. Concentration of biotin-binding sites in chicken egg yolk determined by radioligand exchange at 45 °C (O) and 65 °C (●) as a function of dietary biotin
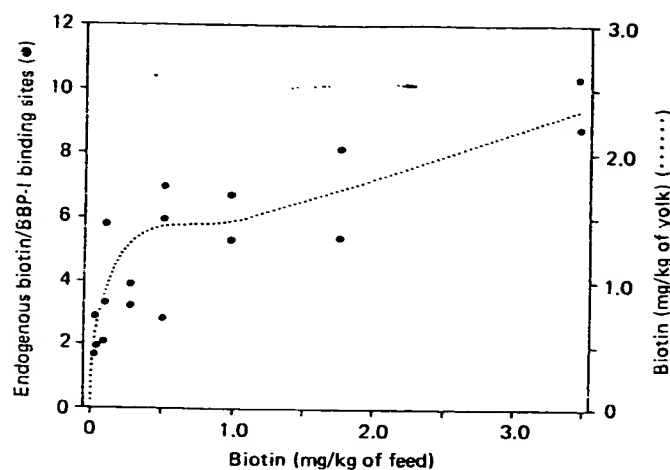
Fig. 8. Estimation of the ratio of endogenous biotin to BBP-1 binding sites in yolk as a function of dietary biotin

The dotted reference line is the microbiologically determined endogenous biotin content of the same samples from Fig. 2 scaled to coincide with the data points.

Table 3. Estimation of the amounts of various biotin-binding proteins in the plasma of immature and adult chickens

Values are expressed as the concentration of biotin-binding sites as determined by radioligand exchange assays for 40 min at 45 °C for BBP-II, 65 °C for BBP-I and 85 °C for avidin.

		Concn. (nM)		
Age	Sex	BBP-II	BBP-I	Avidin
1 day	Male	< 0.6	~ 1.8	< 1.0
	Female	< 0.5	~ 2.3	< 1.5
2 weeks	Male	< 1.0	~ 2.6	< 1.3
3 weeks	Female	< 0.5	~ 1.5	< 0.5
6 weeks	Female	< 0.8	~ 3.0	< 0.8
8 weeks	Male	< 1.0	~ 4.2	< 0.9
9 weeks	Female	< 0.8	~ 2.7	< 0.7
11 weeks	Male	< 0.6	~ 2.3	< 1.2
Adult	Female	40–120	25–40	n.d.*
	Male	~ 0	< 2.9	10.5

* n.d., not determined.

significant amounts only in the plasma of adult males. In all samples, the amount of biotin equals or exceeds the available binding sites.

DISCUSSION

Discovery of BBP-II

As originally designed, there were two objectives in the present study. The first was to determine if biotin transport to the yolk was limited by BBP. The other was to determine if apo-BBP could be generated *in vivo* and be transported to yolk. At the time, only one BBP was known from yolk and it was stable at 65 °C. Our initial results at 65 °C (Fig. 8) with a new assay based on [³H]biotin exchange produced the expected saturation profile, but unexpectedly revealed severalfold more biotin in yolk than biotin-binding sites. This paradox was resolved when a second, more abundant but heat-labile, BBP (BBP-II) was discovered. The discovery of BBP-II seems to be dependent in part on the new assay, which uses [³H]biotin of high specific radioactivity and permits a 100-fold higher dilution of samples (L. Bush & H. B. White III, unpublished work). [¹⁴C]Biotin-based analyses of egg yolk which should have contained BBP-II show no heat-labile component below 70 °C (White *et al.*,

1976; Kulomaa *et al.*, 1981). Furthermore, the earlier assay procedure had not been developed so that the presence of excess endogenous biotin could be detected. Although quantification of the amounts of BBP-I and BBP-II in mixtures is imprecise, the general patterns that have emerged are clear, as discussed below.

Relationship of plasma biotin to yolk biotin

The deposition of biotin in yolk is not the linear or saturable function of biotin in the plasma that would be expected for a simple receptor-mediated or diffusional process. Rather, biotin at low concentrations in plasma is transferred to yolk less efficiently than it is at higher concentrations (Fig. 4). This pattern can be qualitatively explained by the presence of two BBPs whose production is differentially dependent on dietary biotin and whose efficiencies of transfer to yolk are different.

The plasma concentration of BBP-I is rather constant in laying hens unless dietary biotin is severely restricted, in which case the concentration of BBP-I is lower. The fact that BBP-I remains saturated with biotin even in biotin-deficient hens (Fig. 8) indicates that the synthesis, secretion or stability of this protein is dependent on biotin. The pattern for the production of BBP-II is similar, except that a greater response is observed and the response saturates at higher dietary biotin. Thus over the

Table 2. Biotin-binding capacity and biotin saturation of avidin in albumen from hens fed different amounts of biotin

Dietary biotin (μg/kg)	28	100	257	557	987	1766	3475
Biotin-binding capacity* (μg/kg of albumen)							
Pretreatment	5.37	7.52	5.34	5.78	5.68	5.25	5.15
Week 1	—	6.49	7.60	5.49	6.15	4.64	4.39
Week 2	—	7.23	5.64	5.53	6.60	4.74	4.88
Week 5	4.66	5.17	4.80	3.86	5.11	4.17	4.31
Biotin saturation (%) (average for weeks 1 and 2)	6.3	22.3	54.3	70.1	81.5	94.4	97.2

* Values reported have been corrected for losses due to avidin binding to surfaces at the high dilutions used in the [³H]biotin-based assay.

normal dietary biotin range the production of BBP-II relative to BBP-I increases from less than 1 at 30 μg of biotin/kg of feed to perhaps greater than 5 at 250 μg/kg. The concentration ratio of BBP-II to BBP-I in yolk is greater than in plasma, indicating that BBP-II is deposited more efficiently in the yolk than is BBP-I.

We conclude that normal biotin deposition in egg is dependent on, and stoichiometric with, BBPs, as previously asserted (White, 1985). However, at very high dietary biotin levels there is increased biotin deposition in yolk in the absence of increased production of BBP. This additional biotin behaves as free biotin in our assay. The unexpectedly high efficiency with which this 'free' biotin is transferred from plasma to yolk suggests that it may in fact be bound weakly to a second site on BBP-II or be associated with lipids (Trager, 1948).

Dietary biotin absorbed in excess of that necessary to saturate BBP-I and BBP-II at maximal production should appear as free biotin; however, as indicated above, 'free' biotin is not detected in plasma or yolk until dietary biotin is considerably increased. This excess biotin in the plasma is scavenged very efficiently by avidin in the oviduct. Only when avidin becomes saturated are there significant increases in free biotin in plasma and yolk (Figs. 1 and 2).

Regulation of the production of BBP-I and BBP-II

The synthesis of both BBP-I and BBP-II is stimulated during egg laying. This implies that the gene (or genes) for these two proteins is (are) induced by oestrogen, as has been suggested by Murty & Adiga (1985). The observation that the production of BBP-I and BBP-II was also dependent on biotin availability was unexpected. Thus both sex hormones and a specific ligand regulate the production of both proteins. The mechanism of the biotin effect is not known. Although transcriptional regulation is possible, a translational regulation can be envisaged in which ligand binding to the nascent protein must occur before termination or secretion. A less efficient biotin-dependent mechanism could depend on the instability of the apoprotein.

Mechanism of biotin deposition in the oocyte

Ultrastructural studies show that yolk deposition occurs via a very active clathrin-mediated endocytosis (Perry et al., 1978, 1984; Griffin et al., 1984). The deposition of specific proteins such as vitellogenin and low-density lipoproteins has been shown to be receptor-mediated (Woods & Roth, 1984; Krummins & Roth, 1981). Despite the presence of these receptors, the concentration of these proteins in yolk is only about six times greater than in plasma. Cholecalciferol is concentrated by this same factor and is thought to be deposited as a complex of its binding protein and the phosvitin moiety of vitellogenin (Fraser & Emtage, 1976). Similarly riboflavin is concentrated 6-fold in yolk (White et al., 1976), yet no receptor has been detected for its binding protein (Benore-Parsons, 1986). Other plasma-derived yolk proteins, such as immunoglobulins, serum albumin and transferrin, are not concentrated in yolk (Schjeide et al., 1976), even though receptors for immunoglobins have been reported.

Biotin is concentrated by more than 20-fold in yolk relative to plasma. This ratio varies from about 3 when chickens are fed biotin-deficient diets to almost 30 when diets of high protein content are fed. In comparison with other yolk-to-plasma concentration ratios, these values are quite high and suggest that a specific receptor-mediated transport system exists for BBP-II and perhaps BBP-I as well.

Transfer of biotin-binding proteins from yolk to chick plasma

Several yolk proteins are transferred to the plasma of the embryo. For instance, the immunity of the hen is transferred to the chick via immunoglobulins deposited in yolk (Loecken & Roth, 1983). Similarly, maternal serum transferrin, in pigeons at least, appears in the chick plasma (Frelinger, 1971). The presence of BBP-I in the plasma of newly hatched chicks (Table 3) suggests that there may be transfer of BBP-I from the yolk. Although this possibility cannot be ruled out, the fact that the concentration of BBP-I remains fairly constant during the rapid growth of a chick implies that most, if not all, BBP-I in older chicks at least is synthesized by the chick and not derived from the yolk. The apparent absence of BBP-II in chick plasma precludes significant yolk-to-plasma transfer.

Relationship of BBP-I to BBP-II

Although it is clear that yolk BBPs are distinct from egg-white avidin (Meslar et al., 1978; Murthy & Adiga, 1984), the structural and genetic relationships between BBP-I and BBP-II are yet to be determined. The fact that their concentrations show different dependences on dietary biotin, that their distribution differs between hen and chick plasma, and that their distributions between soluble and particulate fractions of diluted egg yolk differ (results not shown) suggest that two different gene products are present. However, they could be differently modified products of the same gene. Furthermore, there is the possibility that both are tetrameric isoproteins which can generate hybrid intermediate forms that associate with a receptor or other proteins with different affinities.

Whatever the structural and genetic relationships are between BBP-I and BBP-II, there seems to be a functional differentiation of the two proteins. BBP-I is present in the plasma of immature chickens. Despite its increased concentration in the plasma of laying hens relative to chick plasma, it is transferred to yolk less efficiently than is BBP-II. BBP-II, on the other hand, is detected only in the plasma of laying hens and is normally present at higher concentrations than BBP-I. These patterns suggest that BBP-I has primarily a maintenance function and secondarily it serves to transport biotin to yolk. The primary and perhaps sole function of BBP-II seems to be the transport of biotin to the yolk.

Function of avidin

Avidin has been viewed as one of the many antimicrobial proteins of egg albumen (Tranter & Board, 1982). This view is further supported by the induction of avidin synthesis at the site of tissue injury in chickens (Elo & Korpela, 1984). The absence of significant amounts of biotin in egg albumen suggest that avidin has little role in the biotin nutrition of the embryo. Our results are consistent with a non-nutritive role for avidin and show that biotin deposition in albumen occurs only when chickens are fed on diets that have biotin contents well in excess of those of natural foods. We also observed

a saturated BBP with the stability of avidin in the plasma of adult males (Table 1), confirming the earlier observations of Elo *et al* (1979). Perhaps avidin does have a role in biotin metabolism in cockerels.

Comparison of biotin and riboflavin transport to the chicken egg yolk

Parallel studies on the binding-protein-mediated deposition of riboflavin to the oocyte (White *et al.*, 1986) show a different and much simpler pattern than described here for biotin. Riboflavin availability does not regulate the production of riboflavin-binding protein. Apoprotein is produced and deposited in yolk. From this sample of two vitamins, it is clear that generalizations about the mechanism of protein-mediated vitamin transport to the oocyte may be hard to find.

REFERENCES

Benore-Parsons, M. (1986) Ph.D. Thesis, University of Delaware

Boas, M. A. (1927) Biochem. J. **21**, 712–724

Elo, H. A. & Korpela, J. (1984) Comp. Biochem. Physiol. B **78**, 15–20

Elo, H. A., Kulomaa, M. S. & Tuohimaa, P. J. (1979) Comp. Biochem. Physiol. B **62**, 237–240

Fraser, D. R. & Emtage, J. S. (1976) Biochem. J. **160**, 671–682

Frelinger, J. A. (1971) Science **171**, 1260–1261

Frigg, M. (1976) Poult. Sci. **55**, 2310–2318

Frigg, M. & Brubacher, G. (1976) Int. J. Vitam. Nutr. Res. **46**, 314–321

Frigg, M., Broz, J. & Streiff, K. (1984) Proc. World Poultry Congr. 17th 420–421

Green, N. M. (1975) Adv. Protein Chem. **29**, 85–133

Griffin, H. D., Perry, M. M. & Gilbert, A. B. (1984) in Physiology and Biochemistry of the Domestic Fowl (Bell, D. J. & Freeman, B. M., eds.), vol. 5, pp. 345–380, Academic Press, London

György, P. & Rose, C. S. (1942) Proc. Soc. Exp. Biol. Med. **49**, 294–298

Kögel, F. & Tönnis, B. (1936) Hoppe-Seyler's Z. Physiol. Chem. **242**, 43–74

Krummins, S. A. & Roth, T. F. (1981) Biochem. J. **196**, 481–488

Kulomaa, M. S., Elo, H. A., Niemela, A. O. & Tuohimaa, P. J. (1981) Biochim. Biophys. Acta **670**, 207–213

Loeken, M. R. & Roth, T. F. (1983) Immunology **49**, 21–28

Lotter, S. E., Miller, M. S., Bruch, R. C. & White, H. B., III (1982) Anal. Biochem. **125**, 110–117

Mandella, R. D., Meslar, H. W. & White, H. B., III (1978) Biochem. J. **175**, 629–633

Meslar, H. W. & White, H. B., III (1979) Methods Enzymol. **62**, 316–319

Meslar, H. W., Camper, S. A. & White, H. B., III (1978) J. Biol. Chem. **253**, 6979–6982

Murty, C. V. R. & Adiga, P. R. (1984) Biochim. Biophys. Acta **786**, 222–230

Murty, C. V. R. & Adiga, P. R. (1985) Mol. Cell. Endocrinol. **40**, 79–86

Pennington, D., Snell, E. E. & Eakin, R. E. (1942) J. Am. Chem. Soc. **64**, 469

Perry, M. M., Gilbert, A. B. & Evans, A. J. (1978) J. Anat. **125**, 481–497

Perry, M. M., Griffin, H. D. & Gilbert, A. B. (1984) Exp. Cell Res. **151**, 433–446

Schjeide, O. A., Briles, W. E., Holshouser, S. & Jones, D. G. (1976) Cell Tissue. Res. **166**, 109–116

Trager, W. (1948) J. Biol. Chem. **176**, 133–146

Tranter, H. S. & Board, R. G. (1982) J. Appl. Biochem. **4**, 295–338

Wallace, R. A. (1978) in The Vertebrate Ovary (Jones, R. E., ed.), pp. 469–502, Plenum Publishing Corp., New York

White, H. B., III (1985) Proc. N.Y. Acad. Sci. **447**, 202–211

White, H. B., III & McGahan, T. (1986) Fed. Proc. Fed. Am. Soc. Exp. Biol. **45**, 1653

White, H. B., III, Dennison, B. A., Della Fera, M. A., Whitney, C. J., McGuire, J. C., Meslar, H. W. & Sammelwitz, P. H. (1976) Biochem. J. **157**, 395–400

White, H. B., III, Armstrong, J. & Whitehead, C. C. (1986) Biochem. J. **238**, 671–675

Whitehead, C. C. (1980) Br. J. Nutr. **44**, 151–159

Whitehead, C. C., Armstrong, J. A. & Waddington, D. (1982) Br. J. Nutr. **48**, 81–88

Woods, J. W. & Roth, T. F. (1984) Biochemistry **23**, 5774–5780

Yusko, S., Roth, T. F. & Smith, T. (1981) Biochem. J. **20**, 43–50

[11] Biotinidase

By BARRY WOLF, JEANNE HYMES, and GREGORY S. HEARD

Biotinidase (biotin-amide amidohydrolase, EC 3.5.1.12) hydrolyzes biocytin (ε-N-biotinyllysine) to the vitamin biotin and lysine. Biotinidase does not release biotin that is covalently bound to intact holocarboxylases.[1] Carboxylases must be proteolytically degraded to biocytin or small biotinyl peptides before hydrolysis can occur. Biotinidase is necessary for the endogenous recycling of the vitamin and probably for the liberation of the vitamin from dietary protein-bound sources.[2] Biotinidase activity is deficient in serum and other tissues of most children with the inherited disorder late-onset multiple carboxylase deficiency.[3-5]

In 1954, enzymes in hog liver[6] and in human plasma[7] were described that released biotin from the products of proteolytic degradation of hog liver and from biocytin, respectively. The enzymes appeared to be identical and were called biotinidase or biocytinase. Biotinidase has since been studied in various prokaryotes and in the tissues of many eukaryotes. Some microorganisms can grow only if biotin is included in the medium,[6,8,9] whereas others grow in medium containing either free biotin or biocytin; the latter have biotinidase activity. In animals, the highest specific activity of biotinidase is found in serum.[10,11] Activities are high in liver, kidney, and adrenal glands and low in brain.[11,12] Biotinidase is also

[1] D. V. Craft, N. H. Goss, N. Chandramouli, and H. G. Wood, Biochemistry 24, 2471 (1985).
[2] B. Wolf, G. S. Heard, J. R. Secor McVoy, and H. M. Raetz, J. Inherited Metab. Dis. 7 (Suppl. 2), 121 (1984).
[3] B. Wolf, R. E. Grier, W. D. Parker, S. I. Goodman, and R. J. Allen, N. Engl. J. Med. 308, 161 (1983).
[4] B. Wolf, R. E. Grier, R. J. Allen, S. I. Goodman, and C. L. Kien, Clin. Chim. Acta 131, 272 (1983).
[5] B. Wolf, R. E. Grier, R. J. Allen, S. I. Goodman, C. L. Kien, W. D. Parker, D. M. Howell, and D. L. Hurst, J. Pediatr. (St. Louis) 103, 233 (1983).
[6] R. W. Thoma and W. H. Peterson, J. Biol. Chem. 210, 569 (1954).
[7] L. D. Wright, C. A. Driscoll, and W. P. Boger, Proc. Soc. Exp. Biol. Med. 86, 335 (1954).
[8] L. D. Wright, H. R. Skeggs, and E. L. Cresson, Am. J. Chem. 73, 4144 (1951).
[9] M. Koivusalo, C. Elorriaga, Y. Kaziro, and S. Ochoa, J. Biol. Chem. 238, 1038 (1963).
[10] M. Koivusalo and J. Pispa, Acta Physiol. Scand. 58: 13 (1963).
[11] J. Pispa, Ann. Med. Exp. Biol. Fenn. 43 (Suppl. 5), 1 (1965).
[12] S. F. Suchy, J. R. Secor McVoy, and B. Wolf, Neurology 35, 1510 (1985).

---

ureido ring in the binding process. Biotin binding to streptavidin results in changes in the absorption spectrum, and the tryptophan reactivity of streptavidin is similar to that of avidin.[43] It is possible that for biotinidase and egg-yolk biotin-binding proteins, tryptophan residues in the active center may be required for hydrophobic binding of biotin to these enzymes.

The sequence for streptavidin has recently been published and shows considerable similarity to avidin around the tryptophan residues.[44] Although the sequence for E. coli biotin holocarboxylase synthetase has been determined, it is not known whether this protein requires the tryptophan residues for biotin binding. However, the other biotin-binding proteins (biotinidase, avidin, streptavidin, and two egg-yolk biotin-binding proteins) require tryptophan residues for biotin binding. It is therefore possible that there may be homology in the amino acid sequence around the tryptophan residues in all the known biotin-binding proteins.

Conclusions

The best understood role of biotin in various organisms is its function as the prosthetic group of a small number of enzymes, the biotin-dependent carboxylases. Biotin serves as a covalently bound "CO_2 carrier" in these carboxylases. Apart from these enzymes, biotin holocarboxylase synthetase and biotinidase are enzymes of biotin metabolism. Both enzymes are biotin-binding proteins. Biotin holocarboxylase synthetase has been shown to act at the biotin operon in E. coli and to regulate the synthesis of various enzymes required for biotin synthesis. Biotinidase has been indicated to act as a biotin carrier protein in human serum. Furthermore, considerable progress has been made in understanding the role of egg-yolk biotin-binding proteins in depositing biotin in eggs. Two proteins, BBP1 and BBP2, have been identified in egg yolk and can be induced by biotin.

Most of the biotin-binding proteins appear to require tryptophan residues for biotin binding, and there is considerable homology of the amino acid sequence around certain tryptophans in both avidin and streptavidin. However, the sequences for the other biotin-binding proteins are not known at present. Furthermore, biotin can regulate various enzymes, and the existence of biotin and a biotin-binding protein in nuclei suggests that there might be a transcriptional control of various genes by biotin similar to that known for hormones.

[43] L. Chaiet and F. Wolf, Arch. Biochem. Biophys. 106, 1 (1964).
[44] C. E. Argarana, I. D. Kuntz, S. Birken, R. Axel, and C. R. Cantor, Nucleic Acids Res. 14, 1871 (1986).

present in secretory cells, including fibroblasts and leukocytes, and in pancreatic juice and zymogen granules.[2,13] In general, carnivores and omnivores have higher specific activities of biotinidase in serum than do herbivores; the sheep is an exception.[11]

Biotinidase in blood probably originates from the liver.[14] Biotinidase in serum is sialylated, but in rat liver from which blood is removed by perfusion with physiological saline, biotinidase is asialylated.[13] Results of subcellular fractionation of rat hepatocytes suggest that the enzyme is enriched in the microsomal fraction and, to a lesser extent, in the lysosomal fraction.[11,13] Microsomal biotinidase is localized in the rough endoplasmic reticulum and Golgi apparatus.

Biotinidase has been partially purified from *Streptococcus faecalis*,[9] *Lactobacillus casei*,[15] and from hog kidney,[15] liver,[11] and serum.[11] The enzyme from human plasma[1,16] and serum[17] has been purified to homogeneity.

Assay Methods

Both natural and artificial substrates have been used for determining biotinidase activity in various assays using secondary enzymatic,[18] radiometric,[19,20] colorimetric,[4,15] and fluorometric[21-23] measurement. Some methods require derivatization[21] or chromatographic separation of the reaction products.[19,22]

The simplest and most commonly used method for measuring biotinidase activity uses N-(d-biotinyl) 4-aminobenzoate (BPABA) as substrate. Hydrolysis of BPABA results in the liberation of biotin and 4-aminobenzoate (PABA). We describe two methods based on the use of BPABA. The first method is a modification of the method of Knappe et al.,[15] and is used clinically for the diagnosis of biotinidase deficiency.[4] A procedure,

[13] G. S. Heard, R. E. Grier, D. L. Weiner, J. R. Secor McVoy, and B. Wolf, Ann. N.Y. Acad. Sci. 447, 400 (1985).
[14] D. L. Weiner, R. E. Grier, P. Watkins, G. S. Heard, and B. Wolf, Am. J. Hum. Genet. 34, 56A (1983).
[15] J. Knappe, W. Brummer, and K. Biederbick, Biochem. Z. 338, 599 (1963).
[16] B. Wolf, J. B. Miller, J. Hymes, J. R. Secor McVoy, Y. Ishikawa, and E. Shapira, Clin. Chim. Acta 164, 27 (1987).
[17] J. Chauhan and K. Dakshinamurti, J. Biol. Chem. 261, 4268 (1986).
[18] D. L. Weiner, R. E. Grier, and B. Wolf, J. Inherited Metab. Dis. 8, 101 (1985).
[19] L. P. Thuy, B. Zielinska, L. Sweetman, and W. L. Nyhan, Ann. N.Y. Acad. Sci. 447, 434 (1985).
[20] B. Wolf and J. R. Secor McVoy, Clin. Chim. Acta 135, 275 (1983).
[21] H. Ebrahim and K. Dakshinamurti, Anal. Biochem. 154, 282 (1986).
[22] K. Hayakawa and J. Oizumi, J. Chromatogr. 383, 148 (1986).
[23] H. Wastell, G. Dale, and K. Bartlett, Anal. Biochem. 140, 69 (1984).

based on this method, for semiquantitatively measuring biotinidase activity in blood-soaked filter paper disks has been developed.[24] It is used to screen newborn infants for biotinidase deficiency.[25,26] The second method, that of Wolf and Secor McVoy,[20] is specific and more sensitive. Therefore, it is applicable to leukocytes, fibroblasts, and amniotic cells, which have low activities, and to tissues with high activities, but for which the availability is limited.

Modified Method[15] of Knappe et al.[4]

Principle. Biotinidase is assayed by measuring the hydrolysis of BPABA. The PABA released is diazotized, coupled to a naphthol reagent, and determined by its absorbance at 546 nm.

Because of its limited sensitivity, this method is applicable only for specimens with high activity, including serum, plasma, and homogenates of liver, kidney, and adrenal glands. The presence of compounds other than PABA that contain a free primary aromatic amino group, such as sulfonamides, can result in the nonspecific formation of colored compounds. Therefore, the appropriate controls must be used.

Reagents

Potassium phosphate buffer, 50 mM (pH 6.0)
Trichloroacetic acid solution (TCA), 30% (w/v)
Sodium nitrite solution, 0.1% (w/v), prepared on each day of assay
Ammonium sulfamate solution, 0.5% (w/v)
N-1-Naphthylethylenediamine dihydrochloride solution, 0.1% (w/v); this solution is light sensitive and should be stored in a dark bottle
Buffer A, prepared by dissolving potassium ethylenediaminetetraacetic acid (EDTA) in 50 mM potassium phosphate buffer to a final concentration of 50 mM and adjusting the pH to 6.0
Buffer B, prepared by dissolving BPABA in 0.1-0.2 ml sodium bicarbonate, 1.0 M, and adding to Buffer A to achieve a final concentration of 0.15 mM BPABA
Standard solution, prepared by dissolving 27.42 mg of analytically pure PABA in 100 ml of Buffer A; the solution is refrigerated, and for use it is diluted 10-fold, so that 0.1 ml corresponds to 20 nmol of PABA

[24] G. S. Heard, J. R. Secor McVoy, and B. Wolf, Clin. Chem. (N.Y.) 30, 125 (1984).
[25] B. Wolf, G. S. Heard, L. G. Jefferson, V. K. Proud, W. E. Nance, and K. A. Weissbecker, N. Engl. J. Med. 313, 16 (1985).
[26] G. S. Heard, B. Wolf, L. G. Jefferson, K. A. Weissbecker, W. E. Nance, J. R. Secor McVoy, A. Napolitano, P. L. Mitchell, F. W. Lambert, and A. S. Linyear, J. Pediatr. (St Louis) 108, 40 (1986).

Procedure. One hundred microliters of serum, plasma, or tissue extract is added to 1.9 ml of Buffer B in 12 × 75 mm glass tubes at 37°; once the enzyme source has been added, the tube is vortexed, returned to the water bath, and incubated for 30 min. The following control tubes are also carried through the assay in a final volume of 2.0 ml: (1) buffer blank containing 0.1 ml of buffer A and 1.9 ml of buffer B, (2) PABA standard containing 0.1 ml of standard solution (20 nmol PABA) in 1.9 ml of Buffer B, and (3) sample blank containing 0.1 ml of sample and 1.9 ml of buffer A.

The enzymatic reaction is stopped by adding 0.2 ml of TCA at 4°. The acidified samples are centrifuged at 2000 g for 10 min at 25°. A 1.5-ml aliquot is removed from each sample and control, added to a 12 × 75 mm glass test tube, and diluted with 0.5 ml of water. The PABA in each tube is determined by the following procedure. To each diluted aliquot, 0.2 ml of sodium nitrite solution is added, and the solution is mixed. After 3 min, 0.2 ml of ammonium sulfamate solution is added, and the solutions are mixed. After 3 min, 0.2 ml of N-1-naphthylethylenediamine dihydrochloride solution is added. The solution is mixed, and after 10 min (during which time a purple color develops in all tubes that contain PABA) the optical density (OD) of the solution is read at 546 nm against distilled water. Optical density increases linearly up to 30 nmol of PABA.

Biotinidase activity is calculated according to the following formula:

$$\text{Activity} = \frac{(OD_{sample} - OD_3) \times 20 \times 10}{(OD_2 - OD_1) \times 30}$$

in which the subscripts sample, 1, 2, and 3 signify the optical densities of the sample, buffer blank, standard, and sample blank, respectively; 20 is the number of nanomoles of PABA per tube; 10 is the reciprocal of the sample volume in milliliters; and 30 is the incubation time in minutes. Activity is expressed as PABA released, in nanomoles per minute per milliliter of serum. The biotinidase activity in normal human serum ranges from 4.4 to 10.0 nmol/min/ml serum.[4]

Method of Wolf and Secor McVoy[20]

Principle. Biotinidase is assayed by measuring the hydrolysis of N-(d-biotinyl) [carboxyl-14C]PABA. The [carboxyl-14C]PABA released is determined by liquid scintillation counting after separation from unhydrolyzed radioactive substrate.

Reagents

N-(d-Biotinyl)-4-[carboxyl-14C]aminobenzoic acid solution, 2.3 mCi/mmol, 0.6 mM[27]
Potassium phosphate buffer, 0.1 mM (pH 6.0)
p-Hydroxymercuribenzoate solution, 2 mM
Avidin solution, 1% (w/v)
Bentonite suspension, 4.2% (w/v), in potassium phosphate buffer, 0.1 mM (pH 6.0)
Liquid scintillation counting fluid suitable for aqueous solutes

Procedure. Serum is diluted to 1% (v/v) with phosphate buffer. Extracts of leukocytes,[28] fibroblasts, and other tissues are prepared by homogenizing in phosphate buffer in the presence of Triton X-100.[20] Extracts should contain 2-3 mg protein/ml. One hundred microliters of each diluted serum (1%, v/v), leukocyte extract, or fibroblast extract is added to 18 μl of potassium phosphate buffer in a 1-ml polystyrene centrifuge tube. After preincubation for 5 min at 37° in a water bath, the reaction is started by the addition of 15 μl of N-(d-biotinyl) [carboxyl-14C]PABA to each tube. The solution is mixed, and the tube is returned to the water bath. A sample blank containing 100 μl of enzyme solution, 18 μl of p-hydroxymercuribenzoate solution, and 15 μl of substrate solution is also carried through the assay.

After incubation for 90-240 min,[29] the reactions are stopped by the addition of 18 μl of p-hydroxymercuribenzoate (18 μl of phosphate buffer is added to blanks). Twenty-five microliters of avidin solution is added, and the solution is mixed and then incubated at 25° (ambient) for 30 min. To each tube is added 75 μl of bentonite suspension. The suspension is mixed, incubated for 10 min, centrifuged at 2000 g for 2 min (Fisher Microfuge), and then 125 μl of the supernatant is transferred to a scintillation vial. Eight milliliters of liquid scintillation fluid is added, and radioactivity is determined in a liquid scintillation spectrometer. The activity in

27 Radioactive substrate with a specific activity of 2.3 mCi/mmol may be prepared, as described in Ref. 4, from biotin N-hydroxysuccinimide ester (17 mg) and a mixture of [carboxyl-14C]PABA (0.2 mg, 58.5 mCi/mmol; Research Products International Corp., Mount Prospect, IL) and unlabeled PABA (7.1 mg).
28 B. Wolf and L. E. Rosenberg, J. Clin. Invest. 62, 931 (1978).
29 The duration of the incubation interval over which the release of [carboxyl-14C]PABA increases linearly depends on the source of enzyme. One hundred microliters of diluted serum, leukocyte extract, or fibroblast extract can be incubated satisfactorily for 180, 120, and 300 min, respectively. We have found incubation intervals of 120, 90, and 240 min to be practical for serum, leukocytes, and fibroblasts, respectively.

each sample is calculated according to the following formula:

$$\text{Activity} = \frac{(\text{cpm}_{sample} - \text{cpm}_{blank}) \times V_2}{\text{cpm}_{substrate} \times V_1 \times t \times SA}$$

in which cpm_{sample} and cpm_{blank} are the counts per minute associated with the sample and blank tubes; $\text{cpm}_{substrate}$ is counts per minute per 15 μl of substrate (typically ~3000 cpm); V_1 and V_2 are volumes (milliliters) of the counted aliquot and the total stopped reaction solution, respectively; t is incubation time (minutes); and SA is the specific activity of the radioactive substrate (counts per minute per nanomole) added to each tube.

Physical Properties

Human serum biotinidase migrates as an α_1-protein on serum electrophoresis.[10,13,25] The pI of the enzyme is between 4.0 and 4.2, which reflects the fact that the enzyme is sialylated in serum.[13] After neuraminidase treatment the enzyme migrates to the β region on electrophoresis.[13] The asialylated enzyme has a pI of 5.1–5.2 on isoelectric focusing.[30] The enzyme is asialylated in homogenates of human liver and rat liver and kidney.[30]

Estimates of the molecular weight of biotinidase determined by gel filtration chromatography in different laboratories vary greatly. The molecular weight of the human plasma enzyme is about 76,000 on Sephacryl S-200 chromatography.[1] We obtained similar results for the human plasma enzyme.[16] The molecular weights of the hog liver and serum enzymes are 115,000 on Sephadex G-100 and Sephadex G-200.[11] The human serum enzyme is about 110,000 on both Sephadex G-100 and Sephacryl S-200.[17]

There is better agreement for the molecular weight of biotinidase determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and sedimentation analysis. The enzyme appears to be a single polypeptide, with a molecular weight of about 66,000–76,000.[1,16,17]

Catalytic Properties

The acyl–enzyme appears to be an intermediate in the biotinidase reaction based on the formation of biotinyl hydroxamate when the enzyme is incubated in the presence of high concentrations of hydroxylamine with either BPABA or biocytin as the substrate.[11,15]

Human serum biotinidase has a K_m for BPABA of 10–34 μM[4,17] and a K_m for biocytin of 5–7.8 μM.[10,17] Human plasma biotinidase has a K_m for

30 B. Wolf, unpublished data (1985).

BPABA of 10 μM and a K_m for biocytin of 6.2 μM.[1] Hog serum enzyme has a K_m for BPABA of 55 μM and a K_m for biocytin of 8 μM.[11] *Lactobacillus casei* and *S. faecalis* enzymes have similar K_m values.[9,15] The K_m values do not vary significantly over the range pH 5.0–7.0. The V_{max} of the human serum enzyme does change with pH.[17] The V_{max} for biocytin at acid pH is higher than for BPABA, whereas the V_{max} for BPABA is higher at neutral pH. The V_{max} for each substrate is similar at pH 7.

Specificity

Biotinidase specifically cleaves *d*-biotinylamides and esters.[11,15] The carboxyl group of biotin is linked through an amide bond with the ε-amino group of a lysine in a highly conserved region of the carboxylases (-ala-met-biocytin-met-).[31] Biotinidase does not cleave biotin from intact holocarboxylases.[1,17] The rate of hydrolysis of biotin from natural and synthetic biotinylated peptides decreases with increasing size of the peptide, suggesting that biocytin or small biotinyl peptides are the primary substrates of biotinidase *in vivo*.[1]

N-*d*-BPABA is hydrolyzed by hog liver and serum biotinidase at pH 6.5 and by hog kidney biotinidase at pH 6.0 at the same rate as is biocytin.[11,15] N-*l*-BPABA is cleaved at 4–15% of the rate of the *d* isomer, and its K_m is 100-fold higher.[11,15] *d*-Biotin methyl ester is cleaved by the hog liver and serum enzymes at about half the rate of biocytin, whereas hog kidney biotinidase cleaves the methyl ester slightly more rapidly than biocytin.[11,15]

The specificity of hydrolysis of the biotinyl substrates by biotinidase lies in the ureido portion of the biotin molecule. Removal of the sulfur (to yield dethiobiotin) barely alters the rate of hydrolysis.[11,17] However, oxidation of the sulfur in sulfo-BPABA or shortening of the aliphatic side chain (e.g., in nor-BPABA), decreases the activity to 10 or 2% of that of BPABA, respectively.[11] The amide portion is also important. N-Biotinyl 3-aminobenzoate is cleaved at 10% of the rate of the *para* compound, whereas the *ortho* derivative is not cleaved.[17] Human serum biotinidase is specific for the lysine portion of the substrate and does not cleave compounds such as biotinylphenylalanine, biotinyl-ε-aminocaproic acid or biotinyl-γ-amino-*n*-butyric acid.[17] Biotinidase cannot cleave biocytin that is bound to avidin.[11]

Partially purified biotinidases from *S. faecalis* and *L. casei* have different specificities from those described above.[9,15] Both biotinamide and

31 D. B. Rylatt, D. B. Keech, and J. C. Wallace, *Arch. Biochem. Biophys.* 183, 113 (1977).

BPABA are cleaved at much greater rates than biocytin. Biotinidase from *S. faecalis* can cleave biotinyl-β-alanine and biocytin at approximately equal rates.

Inhibition

Biotin is a competitive inhibitor of biotinidase at acidic pH when either biocytin or BPABA is the substrate. Biotin is not inhibitory with BPABA as the substrate at pH 7.4,[11] and the K_i for biotin is increased markedly with biocytin as substrate at alkaline pH.[17] Biotinidase activity is not competitively inhibited by dl-dethiobiotin, l-biotin, l-lysine, ε-aminocaproic acid, fatty acids, d-norbiotin, d-biotin sulfone, or PABA.[11,17]

Sulfhydryl inhibitors inactivate biotinidase. Hog liver and serum biotinidases are completely inactivated by 0.1 mM p-chloromercuribenzoate,[11] and the enzymes in human plasma[1] and hog kidney[15] are totally inactivated by 10 μM p-chloromercuribenzoate.[11] Monoiodoacetate also inhibits enzyme activity, whereas N-ethylmaleimide or arsenite does not inhibit.[11] Partially purified biotinidases from *L. casei* and *S. faecalis* are not inhibited by 0.1 mM p-chloromercuribenzoate.[9,15]

Evidence for the presence of a serine hydroxyl group in or near the active site of biotinidase is equivocal. Hog kidney biotinidase has only 15% of normal activity when treated with 1 μM diisopropyl fluorophosphate,[15] and the activity of the human plasma enzyme is inhibited 99% by 0.1 μM phenylmethylsulfonyl fluoride.[1] Phenylmethylsulfonyl fluoride does not inhibit human serum biotinidase at 1 mM.[17] Diisopropyl fluorophosphate at 1 mM does not completely inactivate the hog serum enzyme and inhibits hog liver enzyme only about 30%.[11] Partially purified biotinidase from *L. casei* is inhibited 50% by 0.4 mM diisopropyl fluorophosphate.[15]

The K_i for biotin using BPABA as substrate is 100 μM at pH 5.0 and 1.3 mM at pH 6.5 for hog serum biotinidase.[11] The K_i for the hog liver enzyme is 40 μM at pH 5.0 and 0.9 mM at pH 6.5.[11] The K_i for the human serum enzyme is 87 μM at pH 5.5 and 1.3 mM at pH 7.[17] The K_i for biotin for human serum enzyme when biocytin is the substrate is 20 μM at pH 5.5 and 0.5 mM at pH 7.[17] The lack of inhibition by biotin at pH 7.4 is thought to be due to slower dissociation of biotin from the enzyme. The slow turnover of product results in less available, inhibitable enzyme.[11,17]

The effect of pH on the inhibition of biotinidase by biotin appears not to be due to changes in the affinity of the enzyme for the substrate. The K_m values for BPABA and biocytin are virtually unchanged between pH 5.5 and 7.0.[11,17]

Stability

Biotinidase loses activity during freezing and thawing but is stable when stored at −70°.[1,17] After ammonium sulfate precipitation the human plasma biotinidase loses 67% of its initial activity in 9 months.[1] Dialyzing hog serum biotinidase at between pH 5.0 and 6.0 stabilizes the enzyme.[11] Buffers with a salt concentration of at least 0.1 M prevent loss of activity.[17] Human serum biotinidase can be stored at 4° in 0.1 M phosphate buffer (pH 6.0) containing 1 mM 2-mercaptoethanol and 1 mM EDTA for 1 month without loss of activity.[17]

The addition of mercaptoethanol (0.1–0.2 mM) or dithiothreitol (1–2 mM) appears to protect or stabilize the enzyme.[1,11,17,32] Heavy metals, such as copper and zinc, inhibit biotinidase.[11] Therefore, EDTA is included in storage buffers at 1–10 mM to prevent this inhibition.

Biotinidase from hog serum and from hog liver is stable both during the enzyme assay and during preincubation for up to 30 min at 37°.[11] Denaturation of the enzyme occurs after 30 min at 60°.[11] The human serum enzyme is denatured 60% after 15 min at 60° and 100% after 15 min at 70°.[17]

32 K. Hayakawa and J. Oizumi, *Clin. Chim. Acta* 168, 109 (1987).

[12] Monoclonal Antibody to Biotin

By KRISHNAMURTI DAKSHINAMURTI and EDWARD S. RECTOR

Introduction

The high affinity between the glycoprotein avidin and the vitamin biotin has been exploited in various areas of biological research, particularly in affinity cytochemistry.[12] The restrictions of this method have also been recognized.[1] Various cell lines, such as HeLa, baby hamster kidney, and human skin fibroblasts, have been shown to bind and internalize avidin.[3,4] In view of this we have explored the possibility of producing monoclonal antibodies to biotin for use in studies on the cellular functions

1 E. A. Bayer and M. Wilchek, *Methods Biochem. Anal.* 26, 1 (1980).
2 J. Korpela, *J. Med. Biol.* 62, 5 (1984).
3 K. Dakshinamurti and L. E. Chalifour, *J. Cell. Physiol.* 107, 427 (1981).
4 L. E. Chalifour and K. Dakshinamurti, *Biochem. Biophys. Res. Commun.* 104, 1047 (1982).